



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



# ALGORITMA *STEMMING* UNTUK *BASO PELEMBANG* BERBASIS ATURAN TATA BAHASA

## TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat  
Untuk Memperoleh Gelar Sarjana Teknik  
Pada Jurusan Teknik Informatika

Oleh

**FINA DEVIANA**  
**11551202083**



UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU

PEKANBARU

2019

## **LEMBAR PERSETUJUAN**

### **ALGORITMA *STEMMING* UNTUK *BASO PELEMBANG* BERBASIS ATURAN TATA BAHASA**

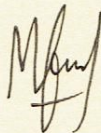
### **TUGAS AKHIR**

Oleh

**FINA DEVIANA**  
**11551202083**

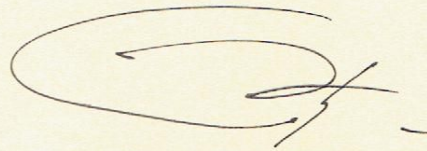
Telah diperiksa dan disetujui sebagai laporan tugas akhir  
di Pekanbaru, pada tanggal 30 Desember 2019

**Pembimbing I,**



**Muhammad Fikry, S.T., M.Sc.**  
**NIP. 19801018 200710 1 002**

**Pembimbing II,**



**Yusra, S.T., M.T.**  
**NIP. 19840123 201503 2 001**



## LEMBAR PENGESAHAN

### ALGORITMA *STEMMING* UNTUK *BASO PELEMBANG* BERBASIS ATURAN TATA BAHASA

### TUGAS AKHIR

Oleh

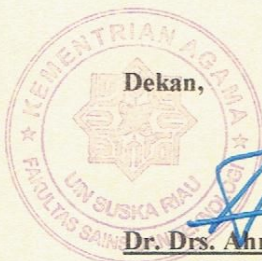
**FINA DEVIANA**  
**11551202083**

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 30 Desember 2019

Pekanbaru, 30 Desember 2019

Mengesahkan,

Ketua Jurusan,



**Dr. Drs. Ahmad Darmawi, M.Ag.**  
**NIP. 19660604 199203 1 004**

**Dr. Elin Haerani, S.T., M.Kom.**  
**NIP. 19810523 200710 2 003**

#### DEWAN PENGUJI

Ketua : Novriyanto, S.T., M.Sc.  
Sekretaris : Muhammad Fikry, S.T., M.Sc.  
Pembimbing II : Yusra, S.T., M.T.  
Anggota I : Elvia Budianita, S.T., M.Cs.  
Anggota II : Suwanto Sanjaya, S.T., M.Kom.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada Penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin Penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya. Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman, dan tanggal pinjam.

Pekanbaru, 30 Desember 2019

**FINA DEVIANA**  
**11551202083**

UIN SUSKA RIAU





**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSEMBAHAN

*Dengan menyebut nama Allah Yang Maha Pengasih lagi Maha Penyayang*

*Alhamdulillah, terima kasihku pada-Mu ya Allah SWT berkat taburan cinta, kasih sayang, rahmat, dan hidayah yang telah Engkau berikan padaku, tugas akhir ini dapat terselesaikan. Shalawat dan salam ku limpahkan untuk Rasulullah Muhammad SAW.*

*Ku persembahkan tugas akhir ini untuk Papa Afrizal Nazir, Mama Mimin, Adik Anneke De Resta, dan Adik Rama Fabian Akbar. Keluargaku tercinta.*

*TERIMA KASIH papa dan mama, atas segala pengorbanan dan perjuangan dalam membesarkanku hingga dapat mencapai pendidikan jenjang S1 ini. Terumtut adik-adikku, semoga karya ini bisa menjadi pedoman dan semangat dalam menyelesaikan pendidikan.*

*~ Fina Deviana ~*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## ALGORITMA *STEMMING* UNTUK *BASO PELEMBANG* BERBASIS ATURAN TATA BAHASA

**FINA DEVIANA**  
**11551202083**

Tanggal Sidang : 30 Desember 2019

Periode Wisuda : aaaaaaa aaaa aa

Jurusan Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau

### ABSTRAK

Bahasa Palembang sehari-hari merupakan perpaduan antara bahasa Melayu dengan bahasa Jawa yang disebut *Baso Palembang*. Pada dokumen *Baso Palembang* ditemukan kata berimbuhan yang tidak terdapat pada kamus. Hal ini menjadi masalah bagi orang awam dalam mempelajari dokumen *Baso Palembang*. *Natural language processing* (NLP) merupakan kemampuan komputer untuk memahami bahasa alami (bahasa manusia) dalam sebuah program. Pembuatan algoritma *stemming* pada tahapan *preprocessing* dalam *text mining* menggunakan NLP digunakan sebagai solusi untuk mengatasi masalah tersebut. Algoritma *stemming* dirancang dengan melakukan penghapusan imbuhan berdasarkan aturan tata bahasa dan pencocokkan kata terhadap *database* kamus sehingga menghasilkan kata dasar. Hasil akurasi sistem berdasarkan kecocokan hasil keluaran dengan kata dasar pada kamus terhadap 645 kata uji dalam *Baso Palembang* adalah 99,84%. Sementara hasil akurasi berdasarkan kecocokan hasil keluaran dengan hasil validasi oleh validator adalah 98,14%. Kegagalan dan kesalahan pada proses *stemming* (*overstemming*) terjadi karena algoritma masih memiliki beberapa ketidaksempurnaan dan ketersediaan kata pada kamus yang masih belum lengkap. Terdapat perbedaan akurasi sekitar 1.7% yang membuktikan bahwa algoritma dapat berjalan dengan baik.

**Kata Kunci:** Algoritma *stemming*, *Baso Palembang*, *Natural language processing*, *Overstemming*, *Text mining*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## STEMMING ALGORITHM FOR BASO PELEMBANG BASED ON GRAMMAR RULES

**FINA DEVIANA**  
**11551202083**

*Date of Final Exam : December, 30<sup>th</sup> 2019*

*Graduation Ceremony Period : 1st, 2nd, 3rd, 4th, 5th*

*Department of Informatics Engineering*

*Faculty of Science and Technology*

*State Islamic University of Sultan Syarif Kasim Riau*

### **ABSTRACT**

*Mother language in Palembang is a combine language of Malay and Javanese that called Baso Palembang. In the Baso Palembang's document, it is found that there are basic words that are not in the dictionary. This is a problem for general people in using the Baso Palembang's document. Natural language processing (NLP) is the ability of computers to understand natural language (human language) in a program. Making the algorithm comes from the preprocessing stage in the text mining using NLP used as a solution to solve this problem. The stemming algorithm was designed by affix removed rules based on grammar and matching words to the dictionary on database. The results of the system accuracy based on the matching output with the basic word in the dictionary against the 645 word test in Baso Palembang is 99.84%. While the results of the accuracy based on the matching output with the results of validation by validator is 98.14%. Failures and errors in the process of stemming (overstemming) occur because the algorithm still has some imperfections and the words in the dictionary is still incomplete. The different accuracy found around 1.7% is proves how the stemming algorithm works well.*

**Kata Kunci:** *Stemming algorithm, Baso Palembang, Natural language processing, Overstemming, Text mining.*



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## KATA PENGANTAR



*Assalamualaikum Warahmatullahi Wabarakatuh*

*Alhamdulillahirabbil'alamin*, segala puji dan rasa syukur atas kehadiran Allah *subhanahu wa ta'ala*, yang telah melimpahkan segala kemudahannya hingga akhirnya Penulis mampu menyelesaikan laporan tugas akhir tepat pada waktunya dengan judul **ALGORITMA STEMMING UNTUK BASO PELEMBANG BERBASIS ATURAN TATA BAHASA**. Laporan tugas akhir ini disusun untuk memenuhi salah satu persyaratan untuk mendapatkan gelar Sarjana Teknik pada Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Selama proses dalam menyelesaikan laporan tugas akhir ini, Penulis telah mendapatkan bantuan, bimbingan, dukungan, serta motivasi, baik secara langsung ataupun tidak langsung. Untuk itu, pada kesempatan ini Penulis ingin menyampaikan ucapan terimakasih kepada:

1. Bapak Prof. Dr. KH. Akhmad Mujahidin, M.Ag., selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Drs. Ahmad Darmawi, M.Ag., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
- Ibu Dr. Elin Haerani, S.T, M.Kom., selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
- Bapak Dr. Alwis Nazir, M.Kom., selaku Pembimbing Akademik dan Ayahanda di kampus yang senantiasa meluangkan waktu untuk membimbing Penulis dalam menjalani perkuliahan di jurusan Teknik Informatika. Memberikan nasehat, motivasi, dan saran dari permasalahan yang dihadapi selama perkuliahan.
- Bapak Muhammad Fikry, S.T., M.Sc., selaku Pembimbing 1 Tugas Akhir yang telah memberikan bimbingan dan semangat kepada Penulis untuk menyelesaikan Tugas Akhir ini.





### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ibu Yusra, S.T., M.T., selaku Pembimbing 2 Tugas Akhir yang telah memberikan bimbingan dan semangat kepada Penulis untuk menyelesaikan Tugas Akhir ini.

Bapak Novriyanto, S.T., M.Sc., selaku Ketua Sidang Tugas Akhir yang telah memberikan motivasi, kritik, dan saran dalam perbaikan laporan ini.

Ibu Elvia Budianita, S.T., M.Cs., selaku Penguji I Tugas Akhir yang telah memberikan motivasi, kritik, dan saran dalam perbaikan dan penulisan laporan ini.

Bapak Suwanto Sanjaya, S.T., M.Kom., selaku Penguji II Tugas Akhir yang telah memberikan motivasi, kritik, dan saran dalam perbaikan dan penulisan laporan ini.

10. Ibu Iis Afrianty, S.T., M.Sc., selaku Koordinator Tugas Akhir yang telah membantu Penulis dalam memperlancar jalannya penyelesaian laporan ini.

11. Seluruh dosen jurusan Teknik Informatika UIN Suska Riau yang telah memberikan banyak ilmu pengetahuan, wawasan, dan pola pikir yang bermanfaat bagi Penulis.

12. Bapak RM Ali Hanafiah (Mang Amin), selaku validator untuk *Baso Pelembang* yang telah membantu Penulis dalam memperlancar jalannya penyelesaian Tugas Akhir ini.

13. Sahabat SDku *Sherwin The Shimmer* (Ayu, Dhea, Dinar, Fera, dan Putri), sahabat SMPku (Ara, Eca, Nana, dan Ndha), sahabat SMAku IS4ADJ terutama (Mamay, Nanat, Sani, dan Winna), Trya Risty, Arifky Nanda Prasetya, Kak Sarah Afrina Sari, Rahma Fitri Asriani, Rizka Hafsari, Tia Sari Indayani, dan Asranita yang telah mau mendengarkan segala drama Tugas Akhir, memberikan dukungan, dan do'a sehingga laporan ini dapat diselesaikan.

14. Teman-teman grup *career center* dan riset teknologi, grup *Awkeren*, grup *Funtastic 4*, kelas F 2015, angkatan 2015, senior, dan junior yang tidak bisa disebutkan satu per satu namanya, terimakasih atas do'a, bantuan, dan dukungannya selama ini sehingga Tugas Akhir Penulis dapat diselesaikan.

15. Semua pihak yang namanya tidak dapat disebutkan satu persatu, yang telah banyak membantu dalam penyelesaian Tugas Akhir ini.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari kesempurnaan. Oleh karena itu, Penulis berharap mendapatkan kritik dan saran yang membangun dari pembaca demi kesempurnaan laporan ini. Kritik dan saran dari pembaca dapat dikirimkan melalui email **Fina.Deviana@students.uin-suska.ac.id** dengan harapan semoga laporan ini dapat memberikan manfaat bagi pembaca. Akhir kata penulis ucapkan terimakasih.

*Wassalammu'alaikum Warahmatullahi Wabarakatuh*

Pekanbaru, 30 Desember 2019

**Penulis**

UIN SUSKA RIAU



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN .....	ii
LEMBAR PENGESAHAN .....	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERSEMBAHAN .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL .....	xvi
DAFTAR LAMPIRAN .....	xvii
DAFTAR PERSAMAAN.....	xviii
DAFTAR SIMBOL .....	xix
BAB I PENDAHULUAN.....	I-1
1.1 Latar Belakang .....	I-1
1.2 Rumusan Masalah .....	I-4
1.3 Batasan Masalah.....	I-4
1.4 Tujuan .....	I-5
1.5 Sistematika Penulisan .....	I-5
BAB II LANDASAN TEORI .....	II-1
2.1 Algoritma .....	II-1
2.2 <i>Natural Language Processing</i> .....	II-1
2.3 <i>Text Mining</i> .....	II-2
2.4 <i>Stemming</i> .....	II-3
2.5 <i>Kesalahan Stemming</i> .....	II-3
2.6 <i>Morfologi Baso Pelembang</i> .....	II-3
2.6.1 Awalan (Prefiks) .....	II-4
2.6.2 Sisipan (Infiks).....	II-8
2.6.3 Akhiran (Sufiks).....	II-9

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.6.4 Kombinasi Imbuhan (Konfiks) .....	II-10
2.7 Rule Based Approach .....	II-13
2.8 Flowchart .....	II-13
2.9 Pseudo Code .....	II-14
2.10 Pengujian White Box .....	II-14
2.11 Pengujian Akurasi Algoritma .....	II-14
2.12 Penelitian Terkait .....	II-15
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>III-1</b>
3.1 Tahapan Penelitian .....	III-1
3.2 Prosedur Penelitian .....	III-2
3.2.1 Rumusan Masalah .....	III-2
3.2.2 Studi Pustaka .....	III-2
3.2.3 Pengumpulan Data .....	III-2
3.2.4 Analisa dan Perancangan .....	III-3
3.2.5 Implementasi dan Pengujian .....	III-3
3.2.6 Kesimpulan dan Saran .....	III-4
<b>BAB IV ANALISA DAN PERANCANGAN .....</b>	<b>IV-1</b>
4.1 Analisa .....	IV-1
4.1.1 Analisa Pengumpulan Data .....	IV-1
4.1.2 Analisa Aturan Penghapusan Imbuhan <i>Baso Pelembang</i> ..	IV-4
4.1.3 Analisa Algoritma <i>Stemming Baso Pelembang</i> .....	IV-8
4.2 Perancangan Algoritma .....	IV-12
4.2.1 Perancangan <i>Flowchart Algoritma Stemming</i> .....	IV-12
4.2.2 Perancangan Fungsi Cek Kamus .....	IV-16
4.2.3 Perancangan Fungsi Cek Prefiks .....	IV-16
4.2.4 Perancangan Fungsi Hapus Prefiks .....	IV-17
4.2.5 Perancangan Fungsi Cek Sufiks .....	IV-18
4.2.6 Perancangan Fungsi Hapus Sufiks .....	IV-18
4.2.7 Perancangan Fungsi Cek Infiks .....	IV-19
4.2.8 Perancangan Fungsi Hapus Infiks .....	IV-20
4.2.9 Perancangan Fungsi <i>Stemming</i> .....	IV-20



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.3 Perancangan <i>Database</i> .....	IV-21
<b>BAB V IMPLEMENTASI DAN PENGUJIAN.....</b>	<b>V-1</b>
5.1 Implementasi .....	V-1
5.2 Komponen Pendukung Implementasi .....	V-1
5.3 Implementasi Algoritma <i>Stemming</i> .....	V-1
5.4 Pengujian.....	V-4
5.4.1 Pengujian <i>White Box</i> .....	V-4
5.4.2 Pengujian Akurasi .....	V-17
5.4.3 Hasil Analisa Pengujian .....	V-19
5.4.4 Kesimpulan Pengujian .....	V-21
<b>BAB VI KESIMPULAN.....</b>	<b>VI-1</b>
6.1 Kesimpulan .....	VI-1
6.2 Saran.....	VI-1
<b>DAFTAR PUSTAKA .....</b>	<b>xcvi</b>
<b>LAMPIRAN A.....</b>	<b>A-1</b>
<b>LAMPIRAN B .....</b>	<b>B-1</b>
<b>LAMPIRAN C .....</b>	<b>C-1</b>
<b>LAMPIRAN D.....</b>	<b>D-1</b>
<b>LAMPIRAN E .....</b>	<b>E-1</b>
<b>LAMPIRAN F .....</b>	<b>F-1</b>

## DAFTAR GAMBAR

Gambar	Halaman
1 Tahapan Preprocessing dalam Text Mining.....	II-2
1 Tahapan Penelitian .....	III-1
1 Proses Pemasukan Data pada Ms.Excel.....	IV-1
2 Proses Penyimpanan Data pada Ms.Excel .....	IV-3
3 Proses Pengurutan Kata Uji .....	IV-3
4 Proses Penghapusan Duplikat Kata Uji .....	IV-4
5 Flowchart Algoritma Stemming Baso Pelembang.....	IV-13
6 Pseudocode Fungsi Cek Kamus .....	IV-16
7 Pseudocode Fungsi Cek Prefiks.....	IV-17
8 Pseudocode Fungsi Hapus Prefiks .....	IV-17
9 Pseudocode Fungsi Cek Sufiks .....	IV-18
10 Pseudocode Fungsi Hapus Sufiks .....	IV-19
11 Pseudocode Fungsi Cek Infiks.....	IV-19
12 Pseudocode Fungsi Hapus Infiks .....	IV-20
13 Pseudocode Fungsi Stemming .....	IV-21
14 Perancangan Database.....	IV-22
15 Perancangan Tabel Kata Dasar .....	IV-23
16 Field untuk Tabel Kata Dasar .....	IV-23
17 Perancangan Tabel Kata Uji .....	IV-24
18 Field untuk Tabel Kata Uji.....	IV-24
19 Memasukkan Data Kata Dasar pada Database.....	IV-25
20 Memasukkan Data Kata Uji pada Database.....	IV-25
1 Proses Pengecekan Kata pada Database .....	V-2
2 Implementasi Memasukkan Kata diampirke.....	V-2
3 Hasil Stemming Kata diampirke .....	V-3
4 Pengujian White Box Pada Kata maco .....	V-4
5 Pengujian White Box Pada Kata bember .....	V-5
6 Pengujian White Box Pada Kata tangkat .....	V-6
7 Pengujian White Box Pada Kata dapus.....	V-7

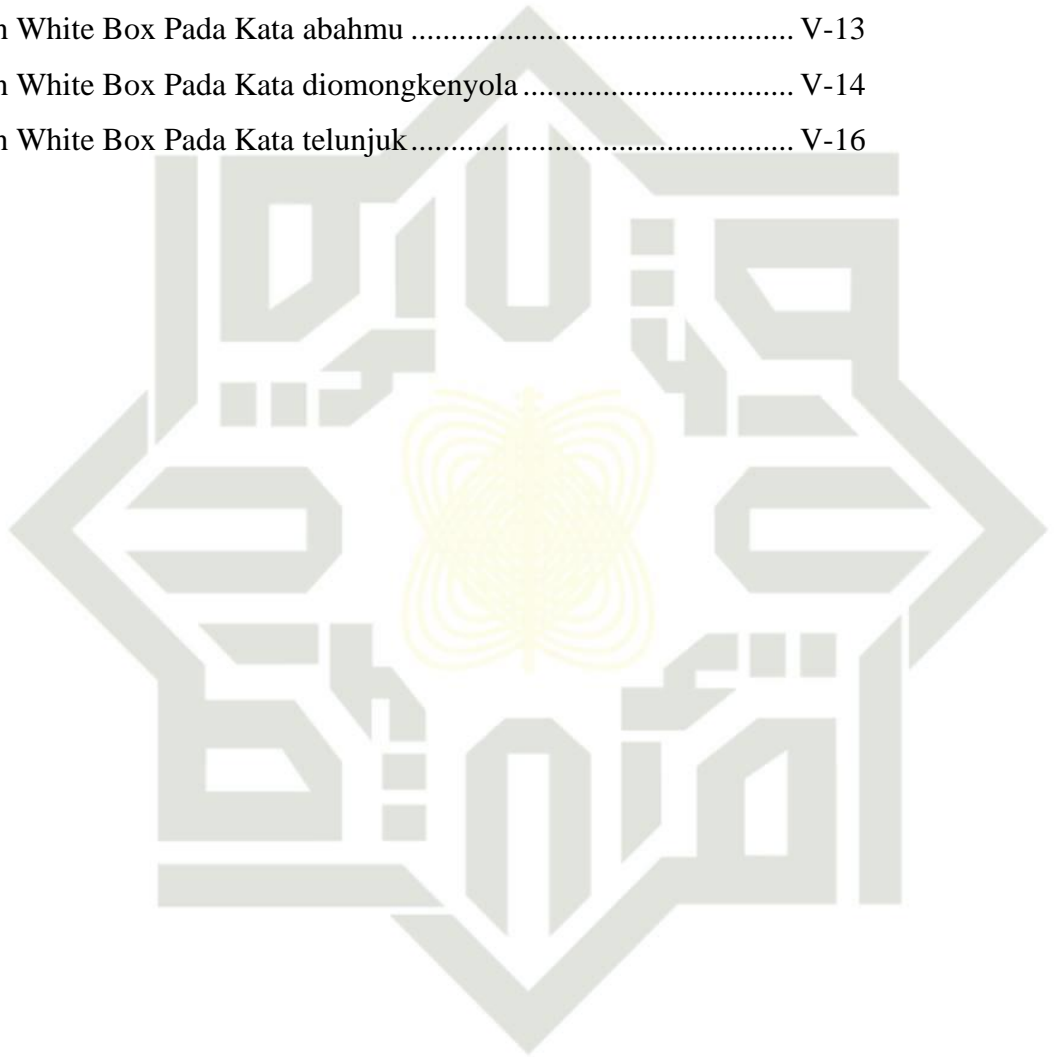




### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5-8	Pengujian White Box Pada Kata siris .....	V-8
5-9	Pengujian White Box Pada Kata penaseat .....	V-8
5-10	Pengujian White Box Pada Kata sepenengeran .....	V-9
5-11	Pengujian White Box Pada Kata gantungke .....	V-11
5-12	Pengujian White Box Pada Kata kujadike .....	V-12
5-13	Pengujian White Box Pada Kata abahmu .....	V-13
5-14	Pengujian White Box Pada Kata diomongkenyola .....	V-14
5-15	Pengujian White Box Pada Kata telunjuk .....	V-16



UIN SUSKA RIAU

## DAFTAR TABEL

Tabel	Halaman
2.1 Penelitian Terkait .....	II-15
4.1 Pengumpulan Kata Dasar .....	IV-2
4.2 Wujud Prefiks Menjadi Kata Dasar .....	IV-5
4.3 Analisa Prefiks se- dan peN- Pada Konfiks .....	IV-5
4.4 Analisa Prefiks se-, pe-, dan N- Pada Konfiks .....	IV-6
4.5 Analisa Prefiks me-, N-, dan ku- Pada Konfiks .....	IV-6
4.6 Aturan Penghapusan Prefiks .....	IV-7
4.7 Wujud Sufiks Menjadi Kata Dasar .....	IV-7
4.8 Analisa Sufiks -mu dan -ku .....	IV-7
4.9 Analisa Sufiks -ke, -nyo, dan -la Pada Konfiks .....	IV-8
4.10 Aturan Penghapusan Sufiks .....	IV-8
4.11 Wujud Infiks Menjadi Kata Dasar .....	IV-8
4.12 Aturan Penghapusan Infiks .....	IV-8
4.13 Proses Penghapusan Prefiks .....	IV-9
4.14 Proses Penghapusan Sufiks .....	IV-9
4.15 Proses Penghapusan Infiks .....	IV-10
4.16 Proses Penghapusan Konfiks .....	IV-11
4.17 Perancangan Struktur Tabel kata_dasar .....	IV-22
4.18 Perancangan Struktur Tabel kata_uji .....	IV-22
5.1 Kata Uji .....	V-17
5.2 Hasil Pengujian Kombinasi Fungsi Utama .....	V-18
5.3 Daftar Kata Gagal .....	V-19
5.4 Daftar Kata Overstemming .....	V-19
5.5 Daftar Kata yang Bukan Kata Dasar .....	V-21

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR LAMPIRAN

Lampiran	Halaman
A Data Kata Uji dan Hasil Stemming .....	A-1
B Flowchart Algoritma Stemming Baso Pelembang .....	B-1
C Lembar Validasi Data Stemming .....	C-1
D Biodata Validator .....	D-1
E Daftar Riwayat Hidup .....	E-1
F Cover Buku Sumber Data .....	F-1

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR PERSAMAAN

Persamaan	Halaman
1 Perhitungan Pengujian Akurasi.....	II-15



UIN SUSKA RIAU

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

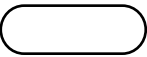
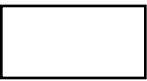
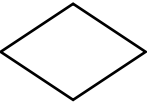


© Hak cipta: milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau



## DAFTAR SIMBOL

Tabel Notasi *Flow Chart*

SIMBOL	NAMA	KETERANGAN
	<i>Terminator</i>	Menyatakan permulaan (awal) atau akhir dari suatu program.
	<i>Process</i>	Menyatakan suatu tindakan (proses) yang dilakukan oleh komputer.
	<i>Decision</i>	Menunjukkan suatu kondisi tertentu yang akan menghasilkan dua kemungkinan jawaban: Ya ( <i>Yes</i> ) / Tidak ( <i>No</i> ).
	<i>Input/Output Data</i>	Menunjukkan masuknya data atau hasil dari suatu proses.
	<i>Flow Line</i>	Menunjukkan arus/aliran dari proses.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Beragam suku, agama, dan bahasa merupakan bentuk dari kesatuan negara Indonesia. Hal ini menjadikan karakteristik negara Indonesia yang unik dan spesifik dari wilayah Sabang sampai Merauke. Bhinneka Tunggal Ika adalah semboyan nasional Indonesia yang memiliki arti berbeda-beda namun tetap satu dengan makna bahwa keberagaman sosial dan budaya dapat membentuk satu kesatuan negara Indonesia. Salah satu karakteristik negara Indonesia adalah beragamnya bahasa di setiap daerah. Bahasa merupakan hal yang utama dalam komponen kebudayaan. Tanpa adanya bahasa, tidak akan terjadi interaksi sosial di masyarakat. Sebagai komponen budaya, menurut (Amin, Amin, Amin, & Tadjuddin, 2009) bahasa termasuk bentuk atau susunan kalimat serta cara penuturan yang dapat menggambarkan watak umum masyarakat pemiliknya.

Menurut (Kemendikbud, 2015) simbol, makna, penggunaan, dan komunikasi merupakan bagian dari bahasa. Simbol adalah makna yang disampaikan dengan pengucapan bunyi atau penulisan tanda. Bagi keperluan berkomunikasi, makna yang terkandung dari simbol-simbol tersebut digunakan untuk berbicara, menyimak, menulis, dan membaca. Seseorang dapat membatasi suatu makna dengan menggunakan bahasa. Bagi kehidupan intelektual, penting untuk menyusun makna dari dokumentasi kosa kata, bahkan makna-makna yang telah terbentuk dapat digunakan dalam konteks dan situasi yang baru. Bahasa adalah penyampaian makna dalam bentuk ide maupun gagasan yang digunakan sebagai alat komunikasi antar manusia. Ragam bahasa adalah corak atau jenis pemakaian bahasa dalam pergaulan hidup. Salah satu jenis ragam bahasa adalah bahasa daerah. Bahasa daerah merupakan bahasa yang digunakan oleh sekelompok masyarakat atau suku bangsa tertentu dalam pergaulan sehari-hari antar sesama mereka (Amin dkk., 2009).



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tata bahasa menurut (Wedhawati, Nurlina, Setiyanto, & Sukesti, 2006) terbentuk dari kata “tata” dan “bahasa”. Tata berarti aturan, kaidah, atau susunan. Sehingga dapat disimpulkan tata bahasa adalah aturan, kaidah, atau susunan dalam pembentukan dan penggunaan sebuah bahasa. Terdapat dua proses pembentukan kata, yakni proses morfologis dan proses nonmorfologis. Morfem menurut (Zaenal & Junaiyah, 2007) adalah makna yang terkandung dalam satuan bahasa terkecil. Proses morfologis adalah proses penghubungan antar morfem satu dengan lainnya yang membentuk kata-kata. Proses nonmorfologis menurut (Yasiroh, 2013) adalah proses yang tidak melalui proses morfologi dalam pembentukan kata. Proses pembentukan kata tidak hanya terjadi di dalam bahasa Indonesia, namun juga di dalam bahasa daerah yang salah satunya adalah bahasa daerah Palembang atau *Baso Pelembang*.

Menurut (Hanafiah, 1995) Palembang merupakan perpaduan kebudayaan Melayu dan Jawa. Bahasa daerah di kalangan masyarakat Palembang disebut dengan *Baso Pelembang*. Secara fisiologis, *Baso Pelembang* berintikan bahasa Melayu, dan banyak dipengaruhi bahasa Jawa, merupakan bahasa Kraton Palembang sejak zaman Kesunanan Palembang, karena pendiri kekuasaan di Palembang yang melepaskan diri dari pengaruh Majapahit, yaitu Ario-damar, anak dari R. Wijaya dengan ibu dari negeri Campa di Cina, dilahirkan dan dibesarkan dalam lingkungan bahasa Jawa. Selain itu, sepanjang perjalanan kekuasaan di Palembang sampai menjadi Kesultanan Palembang Darussalam, sering terjadi perkawinan antar putera Sunan Palembang dengan puteri raja-raja Jawa, seperti Ratu Sinuhun dengan Pangeran Siding Kenayan yang memerintah tahun 1045-1057 H, memperkuat pengaruh bahasa Jawa terhadap *Baso Pelembang* (Amin dkk., 2009).

Penelitian terhadap *Baso Pelembang* di bidang teknologi informasi masih sangat kurang. Bisa diketahui dari sedikitnya hasil penelitian terhadap dokumen-dokumen dalam *Baso Pelembang* yang berbentuk digital, bahkan dalam bentuk cetakan buku juga sulit ditemukan. Hal ini terjadi karena situasi dan kondisi kehidupan masyarakat tradisional mulai terkikis dan berganti menjadi masyarakat *modern*. Proses akulturasi akibat globalisasi yang terjadi begitu cepat membuat

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

masyarakat mengabaikan bahkan meninggalkan budaya dan bahasa daerah masyarakat tersebut. Ketertarikan masyarakat untuk mempelajari budaya dan bahasa daerah menjadi berkurang.

Dalam mempelajari *Baso Pelembang* diperlukan pengetahuan terhadap tata bahasanya selain harus mengetahui arti kata. Pada dasarnya, untuk mempelajari *Baso Pelembang* diperlukan pengetahuan kata dasar dari suatu kalimat. Hal ini karena dari kata dasar tersebut dapat diketahui kedudukan dan arti dari kalimat *Baso Pelembang*. Dalam mencari kata dasar tersebut dibutuhkan keakurasian, sehingga dapat mengurangi kesalahan dalam penentuan kata dasar.

*Natural Language Processing* (NLP) menurut (Sutojo, Mulyanto, & Suhartono, 2011) adalah program yang dibuat untuk memiliki kemampuan dalam memahami bahasa manusia. *Stemming* menurut (Nopiyanti & Sekarwati, 2014) adalah tahapan transformasi kata-kata dalam sebuah dokumen yang berisi teks ke kata dasarnya untuk meningkatkan performa sistem temu kembali atau *information retrieval*. Dalam pengertian umum, *stemming* merupakan proses pencarian kata dasar dengan cara menghilangkan imbuhan (afiks). Ada beberapa jenis imbuhan (afiks) yakni awalan (prefiks), sisipan (infiks), dan akhiran (sufiks) serta gabungan dari awalan dan akhiran (konfiks). Metode *stemming* untuk satu bahasa berbeda dengan bahasa lain, karena metode ini digunakan untuk mengganti bentuk dari sebuah kata yang memiliki imbuhan menjadi kata dasar sesuai dengan struktur tata bahasa (morfologi) yang baik dan benar. Pembuatan model *stemming* ini akan menghasilkan sebuah kata dasar yang sesuai dalam kamus berdasarkan aturan tata bahasa.

Penelitian terkait perbandingan dan pengaruh alternatif *stemming* pada tahun 2012 oleh Deepika Sharma, pada tahun 2016 oleh Alexandra Schofield dan David Mimno, dan pada tahun 2018 oleh I Made Artha Agastya. Penelitian terkait algoritma *stemming* dapat dirancang berdasarkan aturan morfologi oleh Wahyu Hidayat pada tahun 2017 dan oleh Titin Winarti, Djati Kerami, Lussiana Etp, dan Sunny Arief Sudiro pada tahun 2017. Penelitian terkait pendekatan *Dictionary Base Stemming* telah dilakukan oleh Ahmad Fikri Zulfikar pada tahun 2017.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Penelitian terkait untuk algoritma *stemming* pada tahun 2007 oleh Mirna Adriani, Jelita Asian, Bobby Nazief, dan Hugh E. William merupakan salah satu algoritma *stemming* terbaik untuk Bahasa Indonesia. Hal ini diperkuat dengan adanya penelitian oleh Ledy Agusta pada tahun 2009. Pada tahun 2015, Rinci Kembang Hapsari dan Yunus Juli Santoso menggunakan algoritma Nazief & Adriani untuk penelitiannya. Pada tahun 2016, penelitian yang dilakukan oleh Reina Setiawan, Aditya Kurniawan, Widodo Budiharto, Iman Herwidiana Kartowisastro, dan Harjanto Prabowo juga menggunakan algoritma Nazief & Adriani. Selain algoritma Nazief & Adriani terdapat algoritma *stemming* lainnya, seperti penelitian oleh Dini Nopiyanti dan Kemal Ade Sekarwati pada tahun 2014. Penelitian terkait menentukan kata dasar pada suatu teks berbahasa daerah dengan algoritma *stemming* oleh Rakhmad Maulidi tahun 2016.

Penelitian mengenai *stemming* untuk *Baso Pelembang* hingga saat ini belum dilaksanakan. Oleh karena itu, pada penelitian ini akan dilakukan pembuatan algoritma *stemming* untuk *Baso Pelembang*. Penelitian ini diharapkan dapat memelihara *Baso Pelembang* sebagai aset bahasa, mendukung masyarakat luas dalam pembelajaran penggunaan kata pada *Baso Pelembang*, dan membantu orang awam untuk mengetahui kata dasar pada sebuah dokumen dalam *Baso Pelembang*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah pada penelitian ini adalah bagaimana membuat algoritma *stemming* untuk *Baso Pelembang* berbasis aturan tata bahasa.

## 1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

Aturan tata bahasa berdasarkan buku Morfologi dan sintaksis Bahasa Melayu Palembang tahun 1987 dan buku Tata Bahasa dan Kamus *Baso Pelembang* tahun 2009.

Kamus yang digunakan sebagai sumber data untuk kata dasar berdasarkan buku Tata Bahasa dan Kamus *Baso Pelembang* tahun 2009.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Dokumen teks bahasa daerah yang digunakan sebagai sumber data untuk kata uji berdasarkan novel *Juaro* tahun 2005 dan cerita rakyat pada buku Morfologi dan Sintaksis Bahasa Melayu Palembang tahun 1987.

Aturan tata bahasa berdasarkan proses afiksasi pada morfologi.

## 1.4 Tujuan

Penelitian ini memiliki tujuan sebagai berikut:

Membuat algoritma *stemming* berbasis aturan tata bahasa untuk menemukan kata dasar pada dokumen berbahasa Palembang.

Melakukan pengujian akurasi pada algoritma *stemming* untuk *Baso Pelembang* berbasis aturan tata bahasa yang digunakan untuk menemukan kata dasar *Baso Pelembang*.

## 1.5 Sistematika Penulisan

Agar lebih sistematis dan terarah dalam penyusunan laporan tugas akhir, penulis membagi dalam beberapa bab. Berikut ini adalah bab-bab yang disajikan:

### BAB I PENDAHULUAN

Bab ini berisi deskripsi umum dari penelitian tugas akhir meliputi latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan pada laporan tugas akhir.

### BAB II LANDASAN TEORI

Bab ini berisi tentang teori-teori terkait penelitian tugas akhir meliputi teori algoritma, *natural language processing*, *stemming*, kesalahan *stemming*, *rule based approach*, morfologi *Baso Pelembang*, akurasi algoritma, dan penelitian terkait.

### BAB III METODOLOGI PENELITIAN

Bab ini berisi tentang penjelasan tahapan metode penelitian dan prosedur penelitian yang meliputi rumusan masalah, studi pustaka, pengumpulan data, analisa dan perancangan, implementasi dan pengujian, hingga tahap kesimpulan dan saran.

### BAB IV ANALISA DAN PERANCANGAN

Bab ini berisi tentang analisa yang terdiri dari analisa pengumpulan data, analisa aturan penghapusan imbuhan *Baso Pelembang*, dan



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

analisa algoritma *stemming* serta berisi tentang perancangan algoritma yang terdiri dari perancangan *flowchart* dan fungsi-fungsi algoritma dalam bentuk *pseudo code*, serta perancangan *database*.

### IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi tentang implementasi, komponen pendukung, implementasi algoritma *stemming*, serta berisi tentang pengujian yang terdiri dari pengujian *white box*, pengujian akurasi, hasil analisa pengujian, serta kesimpulan pengujian.

### PENUTUP

Bab ini berisi tentang kesimpulan dan saran dari hasil penelitian tugas akhir.





**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB II

### LANDASAN TEORI

#### 2.1 Algoritma

Menurut (Munir, 2011) algoritma adalah urutan langkah-langkah untuk memecahkan suatu masalah. Terdapat beberapa definisi lain dari algoritma (tetapi pada prinsipnya senada dengan definisi yang diungkapkan sebelumnya) yang dikutip dari beberapa literatur, antara lain:

1. Algoritma adalah deretan langkah-langkah komputasi yang mentransformasikan data masukan menjadi keluaran (Cormen, Leiserson, & Rivest, 1990).
2. Algoritma adalah deretan instruksi yang jelas untuk memecahkan masalah, yaitu untuk memperoleh keluaran yang diinginkan dari suatu masukan dalam jumlah waktu yang terbatas (Levitin, 2003).
3. Algoritma adalah prosedur komputasi yang terdefinisi dengan baik yang menggunakan beberapa nilai sebagai masukan dan menghasilkan beberapa nilai yang disebut keluaran. Jadi, algoritma adalah deretan langkah komputasi yang mentransformasikan masukan menjadi keluaran (Cormen, 1989).

#### 2.2 *Natural Language Processing*

*Natural Language* (bahasa alami) adalah bahasa yang telah berevolusi dan digunakan secara alami oleh manusia untuk tujuan komunikasi. *Natural Language Processing* (NLP) adalah bidang ilmu komputer dan linguistik yang berhubungan dengan interaksi antara komputer dan bahasa (alami) manusia. Istilah bahasa alami digunakan untuk membedakan bahasa manusia (seperti Bahasa Indonesia) dari bahasa formal atau bahasa komputer (seperti *Java*) (Kumar, 2011).

*Natural Language Processing* (NLP) atau dalam bahasa Indonesia merupakan Pengolahan Bahasa Alami yaitu program yang dibuat untuk memiliki kemampuan dalam memahami bahasa manusia. Pada prinsipnya, menurut (Sutojo dkk., 2011) bahasa alami adalah suatu bentuk representasi dari suatu pesan yang ingin

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

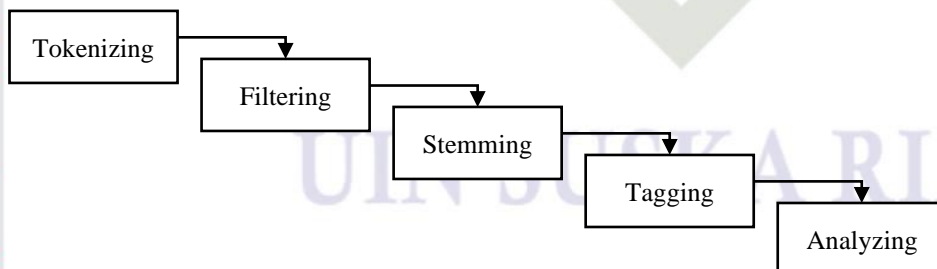
dikomunikasikan antarmanusia. Bentuk utama dalam representasi tersebut berupa suara/ucapan (*spoken language*), tetapi dapat juga dinyatakan dalam bentuk tulisan.

### 2.3 Text Mining

*Text mining* menurut (Sulhan & Kurniawan, 2014) adalah proses semi otomatis penggalian pola (pengetahuan dan informasi yang berguna) dari sejumlah sumber data berukuran besar yang tidak terstruktur berupa teks. *Text mining* menggunakan *natural language processing* untuk memasukkan struktur ke dalam kumpulan teks. *Natural language processing* berupaya memecahkan masalah untuk memahami bahasa alami manusia dan mengubah bahasa tersebut menjadi representasi formal yang dapat diproses oleh komputer.

Sulhan dan Kurniawan menambahkan tahapan *preprocessing* dalam *text mining* secara umum sebagai berikut:

- a. *Tokenizing*, merupakan tahapan pemotongan *string input* berdasarkan tiap kata yang menyusunnya.
- b. *Filtering*, merupakan tahapan mengambil kata-kata penting. Dapat menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *word list* (menyimpan kata penting).
- c. *Stemming*, merupakan tahapan mencari *root* kata.
- d. *Tagging*, merupakan tahapan mencari bentuk awal/*root* dari tiap kata lampau.
- e. *Analyzing*, merupakan tahapan penentuan seberapa jauh keterhubungan antar kata-kata antar dokumen yang ada.



**Gambar 2. 1 Tahapan *Preprocessing* dalam *Text Mining***



## 2.4 Stemming

*Stemming* menurut (Nopiyanti & Sekarwati, 2014) adalah tahapan transformasi kata-kata dalam sebuah dokumen yang berisi teks ke kata dasarnya. Dalam pengertian umum, *stemming* merupakan proses pencarian kata dasar dengan cara menghilangkan imbuhan (afiks). Ada beberapa jenis imbuhan (afiks) yakni: awalan (prefiks), sisipan (infiks), dan akhiran (sufiks) serta gabungan dari awalan dan akhiran (konfiks).

Suatu proses menurut (Agusta, 2009) yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu disebut *stemming*. Contoh kata mencari, nyari, mencarikan, akan di *stem* ke *root word* menjadi kata asal yaitu “cari”.

## 2.5 Kesalahan Stemming

Menurut (Paice, 1996), terdapat dua jenis kesalahan dalam *stemming*, yaitu:

1. *Understemming*, adalah proses *stemming* yang menghasilkan kata yang tidak terkelompok dalam satu kelompok atau kelas. Hal ini menyebabkan konsep tunggal yang tersebar di beberapa hasil *stem* yang berbeda. Contoh dalam bahasa Indonesia, kata mengecek = meng-ecek yang seharusnya menge-cek.
2. *Overstemming*, adalah proses *stemming* yang menghasilkan kata namun kata tersebut seharusnya tidak diletakan dalam kelompok atau kelas tertentu. Pada proses *stemming* ternyata kata tersebut dimasukan dalam kelompok lainnya. Contoh dalam bahasa Indonesia, kata berikan = ber-i-kan.

Algoritma *stemming* yang lebih banyak menghasilkan *understemming* dibandingkan *overstemming* dikategorikan dalam algoritma *light stemming*, Sedangkan algoritma *stemming* yang lebih banyak menghasilkan *overstemming* dibandingkan *understemming* dikategorikan dalam algoritma *heavy stemming*.

## 2.6 Morfologi Baso Pelembang

Morfem menurut (Aliana, Nursato, Arifin, Soetopo, & Waif, 1987) adalah satuan bahasa terkecil yang mengandung arti. Morfologi ialah bagian dari ilmu bahasa yang mempelajari bentuk dan kejadian kata-kata (Amin dkk., 2009). Kata *ngembek* (mengambil) dan *meli* (membeli) dibentuk dengan membubuhkan awalan



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

N- pada bentuk dasar *embek* (ambil) dan *beli* (beli). Wujud *N-* pada kedua kata itu berbeda-beda, yaitu *ng-* dalam kata *ngembek* dan *m-* dalam kata *meli*. Variasi wujud dari awalan *N-* ini merupakan peristiwa morfofonemik.

Peristiwa morfofonemik dapat muncul melalui proses afiksasi. Peristiwa morfofonemik melalui proses afiksasi dapat ditinjau dengan mengkaji semua imbuhan dalam penggabungannya dengan kata asal atau bentuk dasar. Imbuhan itu meliputi semua awalan, sisipan, akhiran, dan konfiks. Ditinjau dari pembentukan kata (Aliana dkk., 1987), konstruksi imbuhan (afiks) pada kata dalam *Baso Pelembang*, yaitu sebagai berikut.

#### 2.6.1 Awalan (Prefiks)

Pembentukan imbuhan awalan (prefiks) terdiri dari:

##### a. Awalan N-

Penggabungan awalan *N-* dengan kata dasar muncul dalam berbagai wujud, yaitu *m-*, *n-*, *ny-*, *nge-*, *me-*, dan *ng-*. Contoh pembentukan kata berimbuhan untuk prefiks *N-* dapat dilihat pada tabel 2.1 berikut ini.

**Tabel 2. 1 Contoh Pembentukan Prefiks N-**

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>N-</i>	<i>atep</i> (atap)	<i>ngatep</i> (mengatap)
2.	<i>N-</i>	<i>baco</i> (baca)	<i>maco</i> (membaca)
3.	<i>N-</i>	<i>cekel</i> (pegang)	<i>nyekel</i> (memegang)
4.	<i>N-</i>	<i>denger</i> (dengar)	<i>nenger</i> (mendengar)
5.	<i>N-</i>	<i>emberi</i> (emberi)	<i>ngemberi</i> (mengemberi)
6.	<i>N-</i>	<i>gebuk</i> (pukul)	<i>ngebuk</i> (memukul)
7.	<i>N-</i>	<i>iris</i> (iris)	<i>ngiris</i> (mengiris)
8.	<i>N-</i>	<i>jait</i> (jahit)	<i>nyait</i> (menjahit)
9.	<i>N-</i>	<i>kocek</i> (kupas)	<i>ngocek</i> (mengupas)
10.	<i>N-</i>	<i>kolake</i> (permainkan)	<i>ngulake</i> (mempermainkan)
11.	<i>N-</i>	<i>lumpat</i> (lompat)	<i>melumpat</i> (melompat)
12.	<i>N-</i>	<i>libar</i> (lebar)	<i>ngelibar</i> (melebar)
13.	<i>N-</i>	<i>larike</i> (larikan)	<i>ngelarike</i> (melarikan)
14.	<i>N-</i>	<i>masak</i> (masak)	<i>masak</i> (memasak)
15.	<i>N-</i>	<i>namoke</i> (namakan)	<i>namoke</i> (menamakan)
16.	<i>N-</i>	<i>oloki</i> (bujuk)	<i>ngoloki</i> (membujuk)
17.	<i>N-</i>	<i>petek</i> (petik)	<i>metek</i> (memetik)
18.	<i>N-</i>	<i>rumput</i> (rumput)	<i>ngerumput</i> (merumput)
19.	<i>N-</i>	<i>suap</i> (suap)	<i>nyuap</i> (menyuap)
20.	<i>N-</i>	<i>tulis</i> (tulis)	<i>nulis</i> (menulis)

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Awalan	Kata Dasar	Kata Imbuhan
21.	<i>N-</i>	<i>ulang</i> (ulang)	<i>ngulang</i> (mengulang)
22.	<i>N-</i>	<i>wariske</i> (mewariskan)	<i>ngewariske</i> (mewariskan)
23.	<i>N-</i>	<i>yakini</i> (yakini)	<i>ngeyakini</i> (meyakini)

#### b. Awalan *be-*

Penggabungan awalan *be-* dengan kata dasar muncul dalam tiga wujud, yaitu *ber-*, *be-*, dan *b-*. Contoh pembentukan kata berimbuhan untuk prefiks *be-* dapat dilihat pada tabel 2.2 berikut ini.

**Tabel 2. 2 Contoh Pembentukan Prefiks *be-***

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>be-</i>	<i>anak</i> (anak)	<i>beranak</i> (beranak)
2.	<i>be-</i>	<i>banyu</i> (air)	<i>bebanyu</i> (berair)
3.	<i>be-</i>	<i>campur</i> (campur)	<i>becampur</i> (bercampur)
4.	<i>be-</i>	<i>dagang</i> (dagang)	<i>berdagang</i> (berdagang)
5.	<i>be-</i>	<i>ember</i> (ember)	<i>bember</i> (berember)
6.	<i>be-</i>	<i>gawe</i> (kerja)	<i>begawe</i> (bekerja)
7.	<i>be-</i>	<i>isi</i> (isi)	<i>bisi</i> (berisi)
8.	<i>be-</i>	<i>jalan</i> (jalan)	<i>bejalan</i> (berjalan)
9.	<i>be-</i>	<i>kulit</i> (kulit)	<i>bekulit</i> (berkulit)
10.	<i>be-</i>	<i>lari</i> (lari)	<i>belari</i> (berlari)
11.	<i>be-</i>	<i>mato</i> (mata)	<i>bemato</i> (bermata)
12.	<i>be-</i>	<i>namo</i> (nama)	<i>benamo</i> (bernama)
13.	<i>be-</i>	<i>ongkos</i> (ongkos)	<i>bongkos</i> (berongkos)
14.	<i>be-</i>	<i>puter</i> (putar)	<i>beputer</i> (berputar)
15.	<i>be-</i>	<i>ruma</i> (rumah)	<i>beruma</i> (berumah)
16.	<i>be-</i>	<i>sabon</i> (sabun)	<i>besabon</i> (bersabun)
17.	<i>be-</i>	<i>telok</i> (telur)	<i>betelok</i> (bertelur)
18.	<i>be-</i>	<i>ujan</i> (hujan)	<i>hujan</i> (berhujan)
19.	<i>be-</i>	<i>warung</i> (warung)	<i>bewarung</i> (berwarung)

#### c. Awalan *te-*

Penggabungan awalan *te-* dengan kata dasar muncul dalam dua wujud, yaitu *te-* dan *t-*. Contoh pembentukan kata berimbuhan untuk prefiks *te-* dapat dilihat pada tabel 2.3 berikut ini.

**Tabel 2. 3 Contoh Pembentukan Prefiks *te-***

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>te-</i>	<i>angkat</i> (angkat)	<i>tangkat</i> (terangkat)
2.	<i>te-</i>	<i>baco</i> (baca)	<i>tebaco</i> (terbaca)

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Awalan	Kata Dasar	Kata Imbuhan
3.	<i>te-</i>	<i>cekel</i> (pegang)	<i>tecekel</i> (terpegang)
4.	<i>te-</i>	<i>duduk</i> (duduk)	<i>teduduk</i> (terduduk)
5.	<i>te-</i>	<i>ganti</i> (ganti)	<i>teganti</i> (terganti)
6.	<i>te-</i>	<i>isep</i> (isap)	<i>tisap</i> (terisap)
7.	<i>te-</i>	<i>jual</i> (jual)	<i>tejual</i> (terjual)
8.	<i>te-</i>	<i>kurung</i> (kurung)	<i>tekurung</i> (terkurung)
9.	<i>te-</i>	<i>larang</i> (larang)	<i>telarang</i> (terlarang)
10.	<i>te-</i>	<i>makan</i> (makan)	<i>temakan</i> (termakan)
11.	<i>te-</i>	<i>namo</i> (nama)	<i>tenamo</i> (ternama)
12.	<i>te-</i>	<i>omong</i> (ucap)	<i>tomong</i> (terucap)
13.	<i>te-</i>	<i>pacul</i> (cangkul)	<i>tepacul</i> (tercangkul)
14.	<i>te-</i>	<i>raso</i> (rasa)	<i>teraso</i> (terasa)
15.	<i>te-</i>	<i>sapu</i> (tesapu)	<i>tesapu</i> (tersapu)
16.	<i>te-</i>	<i>tulis</i> (tulis)	<i>tetulis</i> (tertulis)
17.	<i>te-</i>	<i>ulang</i> (ulang)	<i>tulang</i> (terulang)

#### d. Awalan di-

Penggabungan awalan *di-* dengan kata dasar muncul dalam dua wujud, yaitu *di-* dan *d-*. Contoh pembentukan kata berimbuhan untuk prefiks *di-* dapat dilihat pada tabel 2.4 berikut ini.

**Tabel 2. 4 Contoh Pembentukan Prefiks *di-***

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>di-</i>	<i>apus</i> (hapus)	<i>dapus</i> (dihapus)
2.	<i>di-</i>	<i>bela</i> (belah)	<i>dibela</i> (dibelah)
3.	<i>di-</i>	<i>cabut</i> (cabut)	<i>dicabut</i> (dicabut)
4.	<i>di-</i>	<i>dapat</i> (dapat)	<i>didapat</i> (didapat)
5.	<i>di-</i>	<i>endepke</i> (rendahkan)	<i>diendepke</i> (direndahkan)
6.	<i>di-</i>	<i>gunting</i> (gunting)	<i>digunting</i> (digunting)
7.	<i>di-</i>	<i>iris</i> (iris)	<i>diiris</i> (diiris)
8.	<i>di-</i>	<i>jemor</i> (jemur)	<i>dijemor</i> (dijemur)
9.	<i>di-</i>	<i>kirin</i> (kirin)	<i>dikirin</i> (dikirin)
10.	<i>di-</i>	<i>lupoke</i> (lupakan)	<i>dilupoke</i> (dilupakan)
11.	<i>di-</i>	<i>makan</i> (makan)	<i>dimakan</i> (dimakan)
12.	<i>di-</i>	<i>namoke</i> (namakan)	<i>dinamoke</i> (dinamakan)
13.	<i>di-</i>	<i>nyaloke</i> (nyalakan)	<i>dinyaloke</i> (dinyalakan)
14.	<i>di-</i>	<i>oloki</i> (bujuk)	<i>doloki</i> (dibujuk)
15.	<i>di-</i>	<i>pake</i> (pakai)	<i>dipake</i> (dipakai)
16.	<i>di-</i>	<i>rusak</i> (rusak)	<i>dirusak</i> (dirusak)
17.	<i>di-</i>	<i>sapu</i> (sapu)	<i>disapu</i> (disapu)
18.	<i>di-</i>	<i>tijak</i> (pijak)	<i>ditijak</i> (dipijak)
19.	<i>di-</i>	<i>ukur</i> (ukur)	<i>diukur</i> (diukur)



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Awalan	Kata Dasar	Kata Imbuhan
20.	<i>di-</i>	<i>wajibke</i> (wajibkan)	<i>diwajibke</i> (diwajibkan)
21.	<i>di-</i>	<i>yakinke</i> (yakinkan)	<i>diyakinke</i> (diyakinkan)

#### e. Awalan *se-*

Penggabungan awalan *se-* dengan kata dasar muncul dalam dua wujud, yaitu *s-* dan *se-*. Contoh pembentukan kata berimbuhan untuk prefiks *se-* dapat dilihat pada tabel 2.5 berikut ini.

**Tabel 2. 5 Contoh Pembentukan Prefiks *se-***

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>se-</i>	<i>atep</i> (atap)	<i>satep</i> (seatap)
2.	<i>se-</i>	<i>bakul</i> (sebakul)	<i>sebakul</i> (sebakul)
3.	<i>se-</i>	<i>cangker</i> (cangkir)	<i>secangkir</i> (secangkir)
4.	<i>se-</i>	<i>dusun</i> (dusun)	<i>sedusun</i> (sedusun)
5.	<i>se-</i>	<i>encer</i> (encer)	<i>sencer</i> (seencer)
6.	<i>se-</i>	<i>gulung</i> (gulung)	<i>segulung</i> (segulung)
7.	<i>se-</i>	<i>iris</i> (iris)	<i>siris</i> (seiris)
8.	<i>se-</i>	<i>jao</i> (jauh)	<i>sejao</i> (sejauh)
9.	<i>se-</i>	<i>karung</i> (karung)	<i>sekarung</i> (sekarung)
10.	<i>se-</i>	<i>lambat</i> (kasur)	<i>selamat</i> (sekasur)
11.	<i>se-</i>	<i>muda</i> (mudah)	<i>semuda</i> (semudah)
12.	<i>se-</i>	<i>nama</i> (nama)	<i>senama</i> (senama)
13.	<i>se-</i>	<i>panas</i> (panas)	<i>sepanas</i> (sepanas)
14.	<i>se-</i>	<i>ramai</i> (ramai)	<i>seramai</i> (seramai)
15.	<i>se-</i>	<i>sampai</i> (sampai)	<i>sesampai</i> (sesampai)
16.	<i>se-</i>	<i>tahun</i> (tahun)	<i>setahun</i> (setahun)
17.	<i>se-</i>	<i>waktu</i> (waktu)	<i>sewaktu</i> (sewaktu)
18.	<i>se-</i>	<i>yakin</i> (yakin)	<i>seyakin</i> (seyakin)

#### f. Awalan *peN-*

Penggabungan awalan *peN-* dengan kata dasar muncul dalam beberapa wujud, yaitu *peng-*, *pem-*, *peny-*, *pen-*, dan *pe-*. Contoh pembentukan kata berimbuhan untuk prefiks *peN-* dapat dilihat pada tabel 2.6 berikut ini.

**Tabel 2. 6 Contoh Pembentukan Prefiks *peN-***

No	Awalan	Kata Dasar	Kata Imbuhan
1.	<i>peN-</i>	<i>atur</i> (atur)	<i>pengatur</i> (pengatur)
2.	<i>peN-</i>	<i>basu</i> (basuh)	<i>pemasu</i> (pembasuh)
3.	<i>peN-</i>	<i>cukur</i> (cukur)	<i>penyukur</i> (penyukur)
4.	<i>peN-</i>	<i>denger</i> (dengar)	<i>pendenger</i> (pendengar)

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Awalan	Kata Dasar	Kata Imbuhan
5.	peN-	garis (garis)	pengaris (penggaris)
6.	peN-	isap (isap)	pengisap (pengisap)
7.	peN-	jait (jahit)	penjait (penjahit)
8.	peN-	kocek (kupas)	pengocek (pengupas)
9.	peN-	lari (lari)	pelari (pelari)
10.	peN-	mabok (mabuk)	pemabok (pemabuk)
11.	peN-	naseat (nasihat)	penaseat (penasihat)
12.	peN-	pake (pakai)	pemake (pemakai)
13.	peN-	rebus (rebus)	perebus (perebus)
14.	peN-	sapu (sapu)	penyapu (penyakit)
15.	peN-	tiduk (tidur)	peniduk (penidur)
16.	peN-	uji (uji)	penguji (penguji)
17.	peN-	waris (waris)	pewaris (pewaris)
18.	peN-	nyanyi (nyanyi)	penyanyi (penyanyi)

#### 2.6.2 Sisipan (Infiks)

Pembentukan imbuhan sisipan (infiks) terdiri dari:

##### a. Sisipan -el-

Contoh pembentukan kata berimbuhan untuk infiks -el- dapat dilihat pada tabel 2.7 berikut ini.

**Tabel 2. 7 Contoh Pembentukan Infiks -el-**

No	Sisipan	Kata Dasar	Kata Imbuhan
1.	-el-	tunjuk (tunjuk)	telunjuk (telunjuk)

##### b. Sisipan -em-

Contoh pembentukan kata berimbuhan untuk infiks -em- dapat dilihat pada tabel 2.8 berikut ini.

**Tabel 2. 8 Contoh Pembentukan Infiks -em-**

No	Sisipan	Kata Dasar	Kata Imbuhan
1.	-em-	kilau (kilau)	kemilau (kemilau)

##### c. Sisipan -er-

Contoh pembentukan kata berimbuhan untuk infiks -er- dapat dilihat pada tabel 2.9 berikut ini.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 2. 9 Contoh Pembentukan Infiks -er-**

No	Sisipan	Kata Dasar	Kata Imbuhan
1.	-er-	<i>getak</i> (getak)	<i>geretak</i> (geretak)

### 2.6.3 Akhiran (Sufiks)

Pembentukan imbuhan akhiran (sufiks) terdiri dari:

#### a. Akhiran -an

Contoh pembentukan kata berimbuhan untuk sufiks *-an* dapat dilihat pada tabel 2.10 berikut ini.

**Tabel 2. 10 Contoh Pembentukan Sufiks -an**

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>taun</i> (tahun)	<i>-an</i>	<i>taunan</i> (tahunan)

#### b. Akhiran -ke

Contoh pembentukan kata berimbuhan untuk sufiks *-ke* dapat dilihat pada tabel 2.11 berikut ini.

**Tabel 2. 11 Contoh Pembentukan Sufiks -ke**

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>bela</i> (belah)	<i>-ke</i>	<i>belake</i> (belahkan)

#### c. Akhiran -i

Contoh pembentukan kata berimbuhan untuk sufiks *-i* dapat dilihat pada tabel 2.12 berikut ini.

**Tabel 2. 12 Contoh Pembentukan Sufiks -i**

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>racun</i> (racun)	<i>-i</i>	<i>racuni</i> (racuni)

#### d. Akhiran -nyo

Contoh pembentukan kata berimbuhan untuk sufiks *-nyo* dapat dilihat pada tabel 2.13 berikut ini.

**Tabel 2. 13 Contoh Pembentukan Sufiks -nyo**

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>tinggi</i> (tinggi)	<i>-nyo</i>	<i>tingginyo</i> (tingginya)



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

#### e. Akhiran *-la*

Contoh pembentukan kata berimbuhan untuk sufiks *-la* dapat dilihat pada tabel 2.14 berikut ini.

**Tabel 2. 14 Contoh Pembentukan Sufiks *-la***

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>payo</i> (mari)	<i>-la</i>	<i>payola</i> (marilah)

#### f. Akhiran *-lah*

Contoh pembentukan kata berimbuhan untuk sufiks *-lah* dapat dilihat pada tabel 2.15 berikut ini.

**Tabel 2. 15 Contoh Pembentukan Sufiks *-lah***

No	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>dados</i> (jadi)	<i>-lah</i>	<i>dadoslah</i> (jadilah)

### 2.6.4 Kombinasi Imbuhan (Konfiks)

Pembentukan kombinasi imbuhan (konfiks) terdiri dari:

#### a. Konfiks *be-...-an*

Contoh pembentukan kata berimbuhan untuk konfiks *be-...-an* dapat dilihat pada tabel 2.16 berikut ini.

**Tabel 2. 16 Contoh Pembentukan Konfiks *be-...-an***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>be-</i>	<i>pegi</i> (pergi)	<i>-an</i>	<i>bepegian</i> (bepergian)

#### b. Konfiks *ke-...-an*

Contoh pembentukan kata berimbuhan untuk konfiks *ke-...-an* dapat dilihat pada tabel 2.17 berikut ini.

**Tabel 2. 17 Contoh Pembentukan Konfiks *ke-...-an***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>ke-</i>	<i>jao</i> (jauh)	<i>-an</i>	<i>kejaoan</i> (kejauhan)

#### c. Konfiks *peN-...-an*

Contoh pembentukan kata berimbuhan untuk konfiks *peN-...-an* dapat dilihat pada tabel 2.18 berikut ini.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 2. 18 Contoh Pembentukan Konfiks *peN-...-an***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>peN-</i>	<i>idup</i> (hidup)	<i>-an</i>	<i>pengidupan</i> (penghidupan)

#### d. Konfiks N-...-i

Contoh pembentukan kata berimbuhan untuk konfiks *N-...-i* dapat dilihat pada tabel 2.19 berikut ini.

**Tabel 2. 19 Contoh Pembentukan Konfiks *N-...-i***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>N-</i>	<i>idup</i> (hidup)	<i>-i</i>	<i>ngidupi</i> (menghidupi)
2.	<i>N-</i>	<i>gepuk</i> (pukul)	<i>-i</i>	<i>ngepuki</i> (memukuli)
3.	<i>N-</i>	<i>tulis</i> (tulis)	<i>-i</i>	<i>nulisi</i> (menulisi)

#### e. Konfiks N-...-ke

Contoh pembentukan kata berimbuhan untuk konfiks *N-...-ke* dapat dilihat pada tabel 2.20 berikut ini.

**Tabel 2. 20 Contoh Pembentukan Konfiks *N-...-ke***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>N-</i>	<i>puti</i> (putih)	<i>-ke</i>	<i>mutike</i> (memutihkan)
2.	<i>N-</i>	<i>denger</i> (dengar)	<i>-ke</i>	<i>nengerke</i> (mendengarkan)
3.	<i>N-</i>	<i>sapu</i> (sapu)	<i>-ke</i>	<i>nyapuke</i> (menyapukan)

#### f. Konfiks se-...-an

Contoh pembentukan kata berimbuhan untuk konfiks *se-...-an* dapat dilihat pada tabel 2.21 berikut ini.

**Tabel 2. 21 Contoh Pembentukan Konfiks *se-...-an***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>se-</i>	<i>parak</i> (dekat)	<i>-an</i>	<i>separakan</i> (sedekatan)

#### g. Konfiks se- + peN-...-an

Contoh pembentukan kata berimbuhan untuk konfiks *se- + peN-...-an* dapat dilihat pada tabel 2.22 berikut ini.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 2. 22 Contoh Pembentukan Konfiks *se-* + *peN-...-an***

No	Awalan 1	Awalan 2	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>se-</i>	<i>peN-</i>	<i>maen</i> (main)	<i>-an</i>	<i>sepemaenan</i> (sepermainan)
2.	<i>se-</i>	<i>peN-</i>	<i>jingok</i> (lihat)	<i>-an</i>	<i>sepenyingokan</i> (sepenglihatan)

#### h. Konfiks *peN-...-an*

Contoh pembentukan kata berimbuhan untuk konfiks *peN-...-an* dapat dilihat pada tabel 2.23 berikut ini.

**Tabel 2. 23 Contoh Pembentukan Konfiks *peN-...-an***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>peN-</i>	<i>jual</i> (jual)	<i>-an</i>	<i>penjualan</i> (penjualan)
2.	<i>peN-</i>	<i>tulis</i> (tulis)	<i>-an</i>	<i>penulisan</i> (penulisan)
3.	<i>peN-</i>	<i>pake</i> (pakai)	<i>-an</i>	<i>pemakean</i> (pemakaian)
4.	<i>peN-</i>	<i>beli</i> (beli)	<i>-an</i>	<i>pemelian</i> (pembelian)
5.	<i>peN-</i>	<i>ajar</i> (ajar)	<i>-an</i>	<i>pengajaran</i> (pengajaran)

#### i. Konfiks *di-...-ke*

Contoh pembentukan kata berimbuhan untuk konfiks *di-...-ke* dapat dilihat pada tabel 2.24 berikut ini.

**Tabel 2. 24 Contoh Pembentukan Konfiks *di-...-ke***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>di-</i>	<i>jalan</i> (jalan)	<i>-ke</i>	<i>dijalanke</i> (dijalankan)

#### j. Konfiks *di-...-i*

Contoh pembentukan kata berimbuhan untuk konfiks *di-...-i* dapat dilihat pada tabel 2.25 berikut ini.

**Tabel 2. 25 Contoh Pembentukan Konfiks *di-...-i***

No	Awalan	Kata Dasar	Akhiran	Kata Imbuhan
1.	<i>di-</i>	<i>mara</i> (marah)	<i>-i</i>	<i>dimarai</i> (dimarahi)



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2.7 Rule Based Approach

Pendekatan berbasis aturan atau *rule based approach* (Karim, Kaykobad, & Murshed, 2013) adalah salah satu pendekatan *stemming* yang lebih cepat karena pendekatan ini dapat langsung diterapkan selama fase penyusunan kamus dalam pengindeksan dan tidak memerlukan tahap *pre-processing* untuk seluruh corpus. Namun, dalam membuat *stemmer* secara manual dibutuhkan pengetahuan linguistik tentang morfologi suatu bahasa.

## 2.8 Flowchart

*Flowchart* atau bagan alir merupakan metode untuk menggambarkan tahap-tahap penyelesaian masalah (prosedur) beserta aliran data dengan simbol-simbol standar yang mudah dipahami. Tujuan utama penggunaan *flowchart* adalah untuk menyederhanakan rangkaian proses atau prosedur untuk memudahkan pemahaman pengguna terhadap informasi tersebut (Soeherman & Pinontoan, 2008).

Menurut (Munir, 2011) notasi algoritmik dengan bagan alir cocok untuk masalah yang kecil, namun tidak cocok untuk masalah yang besar karena membutuhkan berlembar halaman kertas. Selain itu, pengonversian notasi algoritma ke notasi bahasa pemrograman juga cenderung relatif lebih sukar.

Dalam pembuatan *flowchart* program menurut (Barakbah, Karlita, & Ahsan, 2013) tidak ada rumus atau patokan yang bersifat mutlak, namun demikian terdapat beberapa anjuran yang harus diperhatikan, yaitu:

1. *Flowchart* digambarkan disuatu halaman dimulai dari sisi atas ke bawah dan dari sisi kiri ke kanan.
2. Aktivitas yang digambarkan harus didefinisikan dengan menggunakan bahasa dan simbol yang tepat dan definisi ini harus dapat dimengerti oleh pembacanya.
3. Kapan aktivitas dimulai dan berakhir harus ditentukan secara jelas. Hanya terdapat satu titik awal dan satu titik akhir.
4. Setiap langkah dari aktivitas harus diuraikan dengan menggunakan deskripsi kata kerja, misalkan menghitung nilai rata-rata.
5. Setiap langkah dari aktivitas harus berada pada urutan benar.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. Lingkup dan *range* dari aktivitas yang sedang digambarkan harus ditelusuri dengan hati-hati. Percabangan-percabangan yang memotong aktivitas yang sedang digambarkan tidak perlu digambarkan pada *flowchart* yang sama. Simbol konektor harus digunakan dan percabangannya diletakkan pada halaman yang terpisah atau hilangkan seluruhnya bila percabangannya tidak berkaitan dengan sistem.

7. Gunakan simbol-simbol *flowchart* yang standar.

### 2.9 Pseudo Code

*Pseudo code* (*pseudo* artinya semu atau tidak sebenarnya) adalah notasi yang mirip dengan notasi bahasa pemrograman tingkat tinggi, khususnya (paling sering) bahasa *Pascal* dan *C*. Lebih tepatnya *pseudo code* adalah campuran antara bahasa alami dengan bahasa pemrograman (Munir, 2011).

*Pseudo code* berisi deskripsi dari algoritma pemrograman komputer yang menggunakan struktur sederhana dari beberapa bahasa pemrograman tetapi bahasa tersebut hanya ditujukan agar dapat dibaca manusia. Dalam *pseudo code*, tidak ada sintaks standar yang resmi. Karena itu, *pseudo code* ini dapat diterapkan dalam berbagai bahasa pemrograman. Disarankan untuk menggunakan *keyword* yang umum digunakan seperti: *if*, *then*, *else*, *while*, *do*, *repeat*, *for*, dan lainnya (Barakbah dkk., 2013).

### 2.10 Pengujian White Box

Pengujian *white box* adalah cara pengujian dengan melihat ke dalam modul untuk meneliti kode-kode program yang ada, dan menganalisis apakah ada kesalahan atau tidak. Jika ada modul yang menghasilkan *output* yang tidak sesuai dengan proses bisnis yang dilakukan, maka baris-baris program, variabel, dan parameter yang terlibat pada unit tersebut akan dicek satu persatu dan diperbaiki, kemudian di *compile* ulang (Al Fatta, 2007).

### 2.11 Pengujian Akurasi Algoritma

Pengujian algoritma *stemming* dilakukan dengan menghitung perbandingan jumlah hasil *stemming* kata uji yang benar berdasarkan kamus dengan jumlah

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

seluruh daftar kata uji. Berikut ini perhitungan pengujian akurasi (Maulidi, 2016) yang dilakukan:

$$\text{Akurasi} = \frac{\text{jumlah kata valid stem}}{\text{jumlah kata uji coba}} \times 100\% \dots\dots\dots(2.1)$$

## 2.12 Penelitian Terkait

Penelitian-penelitian yang pernah dilakukan tentang *stemming* pada tabel 2.1 sebagai berikut.

**Tabel 2. 26 Penelitian Terkait**

No	Peneliti	Judul	Abstrak
	(Adriani, Asian, Nazief, & Williams, 2007)	Stemming Indonesian : A Confix Stripping Approach	Penelitian ini melakukan pembuatan <i>stemming</i> pada kata-kata berbahasa Indonesia sesuai morfologinya. Hasil dari penelitian adalah algoritma CS ( <i>Confix Stripping</i> ) yang baru dan sangat akurat serta mengeksplorasi keefektifan <i>stemming</i> dalam konteks <i>information retrieval</i> .
2.	(Agusta, 2009)	Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Andriani Untuk Stemming Dokumen Teks Bahasa Indonesia	Pengukuran efektifitas algoritma <i>stemming</i> berdasarkan pada kecepatan proses, keakuratan, dan kesalahan. Ujicoba perbandingan algoritma menggunakan teks berbahasa Indonesia. Hasil dari penelitian membuktikan bahwa algoritma nazief & adriani lebih tepat untuk tujuan linguistik daripada algoritma porter.
	(Sharma, 2012)	Stemming Algorithms: A Comparative Study and Their Analysis	Ada beberapa cara untuk melakukan <i>stemming</i> mulai dari metode manual hingga metode otomatis, dari bahasa khusus ke bahasa umum yang memiliki keunggulan masing-masing. Penelitian ini merupakan studi perbandingan berbagai alternatif <i>stemming</i> yang tersedia untuk meningkatkan efektifitas dan efisiensi <i>information retrieval</i> .



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti	Judul	Abstrak
4.	(Nopiyanti & Sekarwati, 2014)	Aplikasi Pencarian Kata Dasar Dokumen Berbahasa Indonesia dengan Metode Stemming Porter Menggunakan PHP & MYSQL	<i>Stemming Porter</i> dalam program, memiliki 5 tahapan untuk menghapus imbuhan pada kata. Ujicoba dilakukan untuk menilai keberhasilan aplikasi dalam mencari kumpulan kata pada 20 artikel dokumen <i>Prosiding Kommit</i> 2012. Hasil pengujian adalah terciptanya sebuah aplikasi yang dapat mencari kumpulan kata dasar untuk bahasa Indonesia dan beroperasi dengan baik.
5.	(Hapsari & Santoso, 2015)	Stemming Artikel Berbahasa Indonesia dengan Pendekatan Confix-Stripping	<i>Confix Stripping</i> merupakan algoritma untuk melakukan pemenggalan kata imbuhan awalan, akhiran, dan gabungan antara awalan-akhiran menjadi sebuah kata dasar. Berdasarkan hasil pengujian algoritma <i>confix stripping</i> memiliki rata-rata nilai akurasi sebesar 94.85% dari 20 dokumen teks berbahasa Indonesia yang diuji.
6.	(Setiawan, Kurniawan, Budiharto, Kartowisastro, & Prabowo, 2016)	Flexible Affix Classification for Stemming Indonesian Language	Penelitian ini bereksperimen dengan 1704 dokumen teks dengan 255182 token dan kata-kata yang ter-stem 3648 kata. Pada penelitian ini dibandingkan kinerja pendekatan fleksibilitas <i>affix classification</i> dengan pendekatan <i>confix stripping</i> . Hasil dari penelitian ini dapat menutupi kegagalan kata-kata yang duplikat dari pendekatan <i>confix stripping</i> .
	(Maulidi, 2016)	Stemmer untuk Bahasa Madura dengan Modifikasi Metode Enhanced Confix Stripping Stemmer	<i>Enhanced Confix Stripping</i> merupakan algoritma untuk teks bahasa Indonesia dengan tingkat keakuratan yang tinggi, algoritma ini dimodifikasi pada <i>rule base</i> -nya sesuai morfologi bahasa Madura. Ujicoba algoritma ini menggunakan data uji berupa teks/puisi berbahasa Madura dan dilakukan evaluasi hasil dari tingkat akurasi, <i>precision</i> , <i>recall</i> , dan <i>F-Measure</i> .

No	Peneliti	Judul	Abstrak
9.	(Schofield & Mimno, 2016)	Comparing Apples to Apple : The Effects of Stemmers on Topic Models	Penelitian ini melakukan pelatihan dan evaluasi kumpulan variasi model topik menggunakan beberapa algoritma <i>stemming</i> yang berbeda. Meskipun sering digunakan pemodelan topik, ditemukan bahwa hasil dari <i>stemmer</i> tidak menghasilkan peningkatan dalam kemungkinan dan hubungan, bahkan kenyataannya menurunkan stabilitas topik.
9.	(Hidayat, 2017)	Ekstraksi Kata Dasar Secara Berjenjang ( <i>Incremental Stemming</i> ) Berbasis Aturan Morfologi untuk Teks Berbahasa Indonesia	Metode berjenjang pada penelitian ini berdasarkan urutan tertentu, secara bergantian partikel akhiran dan awalan dilepaskan dari sebuah kata sehingga dihasilkan sebuah kata dasar. Ujicoba metode dilakukan pada 6464 dokumen Al-Qur'an terjemahan Indonesia dengan kamus berisi 5000 kata yang dipilih secara acak dari KBBI. Dari 3432 kata unik yang diproses, 94.7% kata dasar dapat diekstrak secara langsung dan hanya 5.3% yang perlu diproses lebih lanjut. Metode ini dapat memilih kata dasar yang tepat hingga 79.12%.
10.	(Winarti, Kerami, Etp, & Sudiro, 2017)	Improving Stemming Algorithm Using Morphological Rules	<i>Stemming</i> CAT merupakan algoritma <i>stemming</i> baru pada penelitian ini. Hasil kata tidak bergantung pada urutan aturan morfologis. Semua aturan akan diperiksa dan hasil kata disimpan dalam daftar kandidat. Digunakan dua jenis daftar kata yaitu daftar kata yang memiliki lebih dari satu kandidat dan daftar kata dasar sebagai referensi kandidat. Hasil penelitian ini memiliki hasil yang lebih baik, memberikan akurasi yang lebih tinggi daripada sistem yang dikenal.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti	Judul	Abstrak
1.	(Zulfikar, 2017)	Pengembangan Algoritma Stemming Bahasa Indonesia dengan Pendekatan Dictionary Base Stemming untuk Menentukan Kata	Pendekatan <i>Dictionary Base Stemming</i> memberikan solusi untuk melakukan <i>stem</i> kata yang berimbuhan pada bahasa Indonesia dan pengembangan algoritma <i>stemming</i> dalam menentukan kata dasar ( <i>root word</i> ) dengan metode pengembangan sistem OOAD ( <i>Object Oriented Analys Design</i> ).
2.	(Agastya, 2018)	Pengaruh Stemmer Bahasa Indonesia Terhadap Peforma Analisis Sentimen Terjemahan Ulasan Film	<i>Stemmer</i> Sastrawi sebagai algortima yang digunakan dalam penelitian ini. Dataset terdiri dari 5 kategori yakni 100 data, 250 data, 500 data, 750 data, dan 1000 data. Hasil dari penelitian menunjukkan bahwa <i>stemmer</i> tidak memberikan peningkatan akurasi yang stabil. Sehingga <i>stemming</i> dapat mengurangi efisiensi dari analisis sentimen.



**Hak Cipta Dilindungi Undang-Undang**

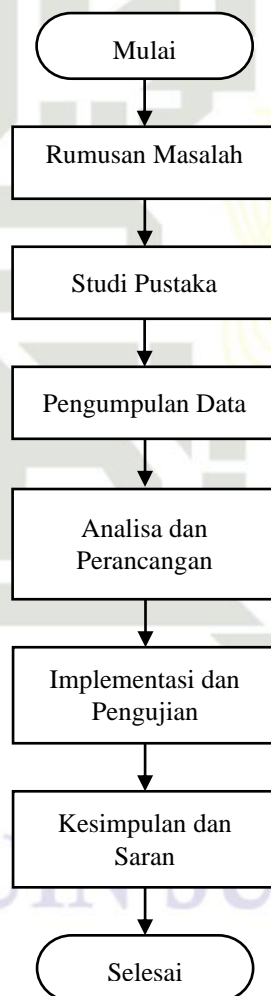
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Tahapan Penelitian

Keberhasilan suatu penelitian ditentukan dari perencanaan penelitian. Dalam perencanaan penelitian tersebut harus jelas apa saja yang akan dilakukan agar penelitian dapat berjalan dengan semestinya. Tahapan-tahapan penelitian ini digambarkan secara umum pada Gambar 3.1 berikut.



**Gambar 3. 1 Tahapan Penelitian**

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 3.2 Prosedur Penelitian

Tahapan penelitian pada Gambar 3.1 dapat dijelaskan sebagai berikut.

### 3.2.1 Rumusan Masalah

Tahapan rumusan masalah pertama kali yang dilakukan adalah menemukan masalah. Setelah masalah ditemukan, dilakukan identifikasi terhadap masalah. Pengidentifikasian masalah dilakukan untuk menghindari pelebaran masalah dengan menyusun batasan-batasan masalah. Setelah batasan masalah tersusun, kemudian dilakukan penyusunan tujuan dan kontribusi penelitian berdasarkan hasil rumusan masalah.

### 3.2.2 Studi Pustaka

Tahapan studi pustaka dilakukan dengan mencari sebanyak-banyaknya sumber penelitian terkait. Sumber penelitian terkait dapat berupa jurnal maupun buku. Hal ini dilakukan untuk menambah wawasan peneliti dalam memahami masalah-masalah yang berkaitan dan pencarian solusi dari masalah tersebut.

### 3.2.3 Pengumpulan Data

Tahapan pengumpulan data dilakukan untuk mendapatkan data yang dibutuhkan dalam penelitian. Ada tiga jenis sumber data yang dibutuhkan dalam penelitian ini, yaitu:

1. Tata Bahasa

Tata bahasa digunakan sebagai sumber data untuk mendukung aturan pembentukan kata dalam penyusunan algoritma *stemming*.

2. Kamus

Kamus digunakan sebagai sumber data untuk pembuatan tabel kata dasar dari bahasa yang akan digunakan pada penelitian. Data kamus diperlukan untuk perbandingan kata pada proses *stemming*. Pada tahapan ini dilakukan pemasukan data kata pada kamus *Baso Pelembang* ke *database* yang telah dibuat. Kegiatan pemasukan data ini untuk mencocokkan kata antara dokumen uji dengan kata pada kamus secara digital.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 3. Dokumen Teks Bahasa Daerah

Dokumen teks bahasa daerah digunakan sebagai sumber data untuk pengujian dari algoritma *stemming* yang dibentuk berdasarkan aturan tata bahasa.

#### 3.2.4 Analisa dan Perancangan

Tahapan analisa dan perancangan dilakukan setelah tahapan pengumpulan data. Tahapan analisa terdiri dari analisa pengumpulan data, analisa aturan penghapusan imbuhan *Baso Pelembang*, dan analisa algoritma *stemming Baso Pelembang*. Analisa pengumpulan data dilakukan untuk mengumpulkan sumber-sumber data yang diperlukan yaitu data kata dasar pada kamus dan data kata uji pada dokumen teks berbahasa daerah. Analisa aturan penghapusan imbuhan *Baso Pelembang* dilakukan untuk menghapus imbuhan yang terdapat pada kata sesuai aturan afiksasi yang terdapat pada buku tata bahasa. Analisa algoritma *stemming Baso Pelembang* dilakukan untuk mencocokkan aturan penghapusan imbuhan dengan algoritma yang akan dibuat pada sistem.

Tahapan perancangan terdiri dari perancangan algoritma dan perancangan *database*. Tahapan perancangan algoritma terdiri dari perancangan *flowchart* algoritma *stemming*, perancangan fungsi cek kamus dalam bentuk *pseudo code*, perancangan fungsi hapus prefiks dalam bentuk *pseudo code*, perancangan fungsi hapus sufiks dalam bentuk *pseudo code*, perancangan fungsi hapus infiks dalam bentuk *pseudo code*, dan perancangan fungsi *stemming* dalam bentuk *pseudo code*. Pada tahapan perancangan *database* dilakukan perancangan tabel –tabel yang akan digunakan pada sistem sebagai *database* yakni tabel untuk kata dasar dan kata uji.

#### 3.2.5 Implementasi dan Pengujian

Tahapan implementasi dan pengujian dilakukan setelah tahapan analisa dan perancangan. Tahapan implementasi terdiri dari implementasi, komponen pendukung implementasi yang terdiri dari komponen perangkat lunak (*software*) dan komponen perangkat keras (*hardware*), serta implementasi algoritma *stemming*.

Tahapan pengujian terdiri dari pengujian *white box*, pengujian akurasi, hasil analisa pengujian, dan kesimpulan pengujian. Pengujian *white box* dilakukan untuk



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

mengetahui kesesuaian kode-kode pada algoritma sistem terhadap proses dengan hasil yang diharapkan. Pengujian akurasi dilakukan untuk mengukur keberhasilan algoritma *stemming* dalam sistem berdasarkan kecocokan hasil keluaran dengan kata dasar yang tersedia pada kamus. Hasil analisa pengujian dilakukan untuk mengetahui kegagalan dan kesalahan yang terjadi pada algoritma *stemming*. Kesimpulan pengujian berisi kesimpulan hasil dari seluruh tahapan pengujian.

**3.2.6 Kesimpulan dan Saran**

Tahapan kesimpulan dan saran dilakukan untuk memberikan penjelasan hasil akhir dari penelitian ini dalam bentuk kesimpulan serta akan disajikan saran-saran untuk penelitian selanjutnya.

UIN SUSKA RIAU

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB IV

### ANALISA DAN PERANCANGAN

#### 4.1 Analisa

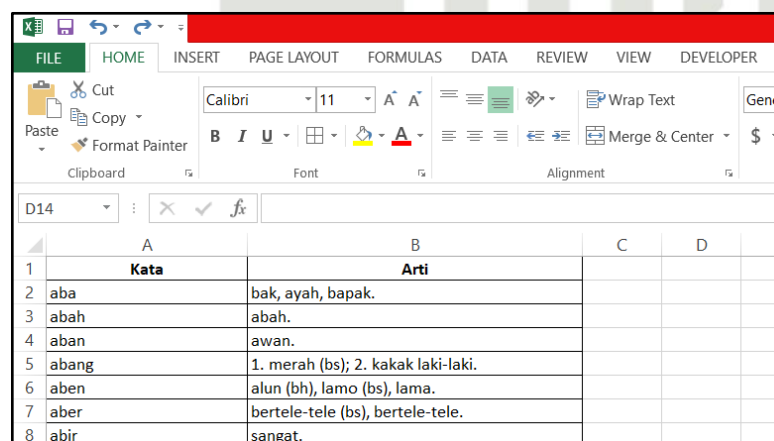
Suatu kemampuan dalam memisahkan konsep secara utuh menjadi beberapa komponen konsep disebut analisa. Terdapat beberapa komponen analisa yang akan dijelaskan yaitu mengenai analisa pengumpulan data, analisa aturan penghapusan imbuhan *Baso Pelembang*, dan analisa algoritma *stemming Baso Pelembang* pada subbab berikut ini.

##### 4.1.1 Analisa Pengumpulan Data

Analisa pengumpulan data terdiri dari pengumpulan kata dasar dan pengumpulan kata uji yang dijelaskan sebagai berikut.

##### 4.1.1.1 Pengumpulan Kata Dasar

Data kamus dalam *Baso Pelembang* yang dikumpulkan bersumber dari buku (Amin dkk., 2009) dan (Aliana dkk., 1987). Fungsi dari data kamus adalah sebagai sumber kata dasar. Proses pengumpulan kata dasar dilakukan secara manual yaitu diketik ulang dengan memasukkan salinan kata yang tersedia di buku ke dalam aplikasi *Microsoft Excel* atau disingkat *Ms.Excel* seperti pada gambar 4.1 berikut.



	A	B	C	D	E
	Kata	Arti			
1					
2	aba	bak, ayah, bapak.			
3	abah	abah.			
4	aban	awan.			
5	abang	1. merah (bs); 2. kakak laki-laki.			
6	aben	alun (bh), lamo (bs), lama.			
7	aber	bertele-tele (bs), bertele-tele.			
8	abir	sangat.			

**Gambar 4. 1 Proses Pemasukan Data pada Ms.Excel**

Kata dasar kemudian disimpan dalam format *Comma Separated Value* (csv) yaitu dengan cara memilih tab *FILE* pilih *Save As* atau tekan tombol *F12* pada

*keyboard* dan pilih lokasi penyimpanan *file*. Kata dasar yang telah dikumpulkan berdasarkan kamus berjumlah 3351 kata yang dijabarkan pada tabel 4.1 berikut ini.

**Tabel 4. 1 Pengumpulan Kata Dasar**

No	Huruf Awal	Jumlah Kata dalam Kamus
1.	a	165
2.	b	346
3.	c	181
4.	d	112
5.	e	51
6.	f	2
7.	g	196
8.	h	14
9.	i	51
10.	j	96
11.	k	379
12.	l	172
13.	m	250
14.	n	300
15.	o	22
16.	p	213
17.	q	0
18.	r	119
19.	s	306
20.	t	286
21.	u	54
22.	v	0
23.	w	28
24.	x	0
25.	y	4
26.	z	4
<b>Total</b>		<b>3351</b>

Selanjutnya pada tahap penyimpanan *file*, nama *file* diisi dengan kata\_dasar pada *File name* dan pilih tipe *file* pada *Save as type* yaitu *CSV (Comma delimited)* seperti pada gambar 4.2 berikut.

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
    - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
  2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

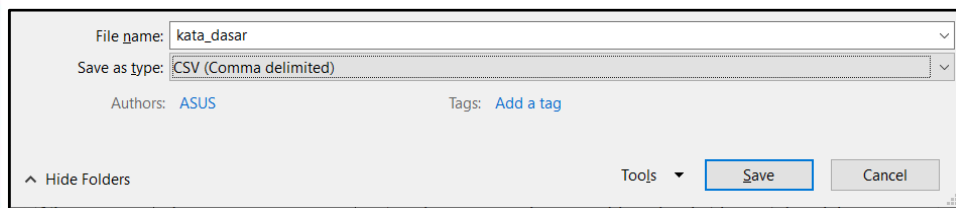


## Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

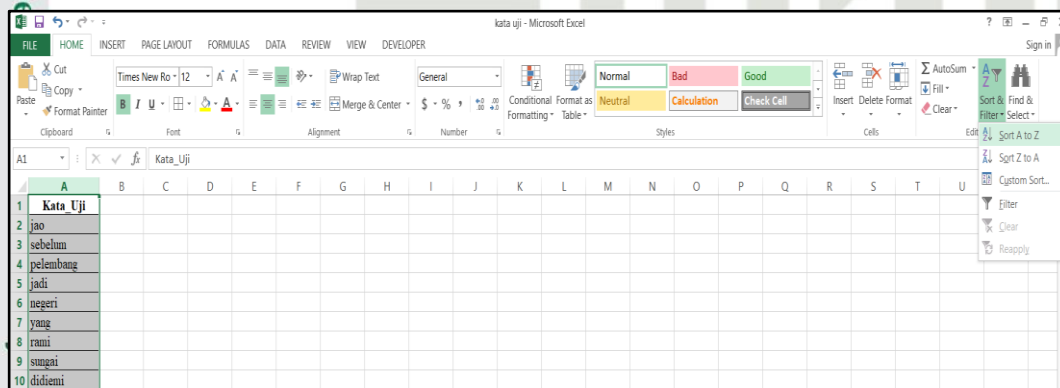
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Gambar 4. 2 Proses Penyimpanan Data pada Ms.Excel**

### 4.1.1.2 Pengumpulan Kata Uji

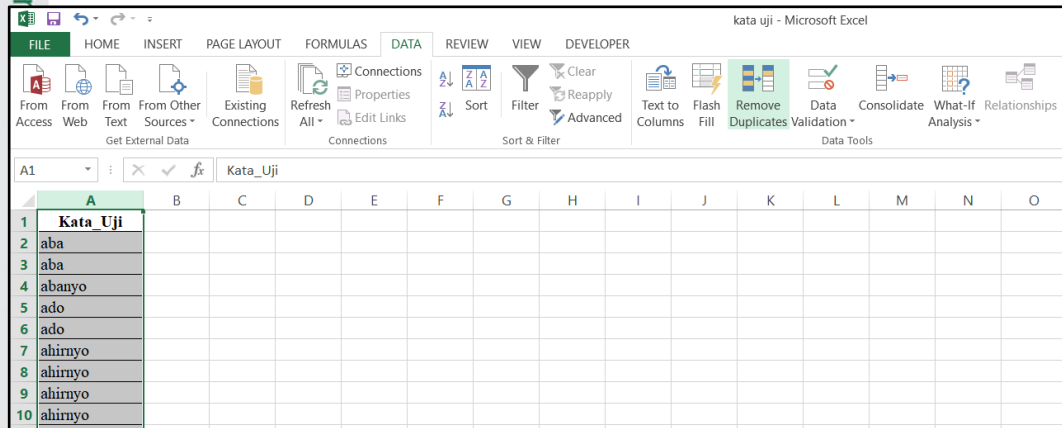
Dokumen berbahasa daerah dalam *Baso Pelembang* yang dikumpulkan bersumber dari buku (Wijaya, 2005) dan (Aliana dkk., 1987). Fungsi dari dokumen berbahasa daerah adalah sebagai sumber kata uji yang terdiri dari kata yang memiliki imbuhan atau kata yang tidak memiliki imbuhan. Proses pengumpulan kata uji dilakukan secara manual sama seperti pengumpulan kata dasar yaitu diketik ulang dengan memasukkan salinan kata yang tersedia di buku ke dalam aplikasi *Microsoft Excel* atau disingkat *Ms.Excel*. Selain itu, kata uji juga dapat dimasukkan satu per satu pada tahap implementasi dan pengujian sistem. Kata uji yang telah dikumpulkan berjumlah 645 kata. Terdapat sebuah kolom yang disediakan yaitu kolom kata\_uji. Sebelum kata uji disimpan dalam bentuk *file*, dilakukan pengurutan kata sesuai abjad. Pengurutan kata dilakukan dengan cara blok kolom kata\_uji, kemudian pada tab *HOME* pilih *Sort & Filter* pada subtab *Editing*. Terdapat beberapa pilihan menu, klik *Sort A to Z* seperti pada gambar 4.3 berikut.



**Gambar 4. 3 Proses Pengurutan Kata Uji**

Setelah kata uji terurut sesuai abjad, kata uji kemudian disimpan dalam format *Comma Separated Value (.csv)* dengan cara yang sama seperti pada tahap penyimpanan kata dasar. Namun sebelumnya, dilakukan penghapusan terhadap

kata uji yang sama (duplikat). Pemeriksaan ini dilakukan dengan cara blok kolom kata\_uji, kemudian pada tab *DATA* klik *Remove Duplicates* dalam subtab *Data Tools* seperti pada gambar 4.4 berikut.



**Gambar 4. 4 Proses Penghapusan Duplikat Kata Uji**

#### 4.1.2 Analisa Aturan Penghapusan Imbuhan *Baso Pelembang*

Kumpulan aturan tata bahasa dalam *Baso Pelembang* yang merupakan susunan pembetuk kata dasar menjadi kata berimbuhan, pada subbab ini akan digunakan sebagai acuan untuk aturan penghapusan kata yang memiliki imbuhan. Hal ini dilakukan sebagai solusi untuk perancangan algoritma *stemming* yang merupakan proses penghapus imbuhan pada kata menjadi kata dasar. Aturan penghapusan imbuhan ini memiliki nomor aturan yang digunakan sebagai pedoman proses kerja penyusunan algoritma. Terdapat 3 kategori aturan penghapusan imbuhan yakni penghapusan prefiks, sufiks, dan infiks yang telah disusun sebagai:

##### a. Aturan Penghapusan Prefiks

Pada aturan penghapusan prefiks akan dilakukan analisa aturan penghapusan imbuhan yang terletak diawal kata. Terdapat 6 prefiks utama yakni *N-*, *be-*, *te-*, *di-*, *se-*, dan *peN-*. Masing-masing dari prefiks ini memiliki beberapa wujud seperti yang dijabarkan pada tabel 4.2 berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 4. 2 Wujud Prefiks Menjadi Kata Dasar**

No	Prefiks Utama	Wujud	Kata Imbuhan	Kata Dasar
1.	<i>N-</i>	a. <i>m-</i> b. <i>me-</i> c. <i>n-</i> d. <i>ny-</i> e. <i>ng-</i> f. <i>nge-</i>	a. <i>maco</i> (membaca) b. <i>melompat</i> (melompat) c. <i>nenger</i> (mendengar) d. <i>nyait</i> (menjahit) e. <i>ngatep</i> (mengatap) f. <i>ngelibar</i> (melebar)	a. <i>baco</i> (baca) b. <i>lumpat</i> (lompat) c. <i>denger</i> (dengar) d. <i>jait</i> (jahit) e. <i>atep</i> (atap) f. <i>libar</i> (lebar)
2.	<i>be-</i>	a. <i>b-</i> b. <i>be-</i> c. <i>ber-</i>	a. <i>bember</i> (berember) b. <i>bebanyu</i> (berair) c. <i>berdagang</i> (berdagang)	a. <i>ember</i> (ember) b. <i>banyu</i> (air) c. <i>dagang</i> (dagang)
3.	<i>te-</i>	a. <i>t-</i> b. <i>te-</i>	a. <i>tangkat</i> (terangkat) b. <i>tecekel</i> (terpegang)	a. <i>angkat</i> (angkat) b. <i>cekel</i> (pegang)
4.	<i>di-</i>	a. <i>d-</i> b. <i>di-</i>	a. <i>dapus</i> (dihapus) b. <i>dicabut</i> (dicabut)	a. <i>apus</i> (hapus) b. <i>cabut</i> (cabut)
5.	<i>se-</i>	a. <i>s-</i> b. <i>se-</i>	a. <i>siris</i> (seiris) b. <i>sebakul</i> (sebakul)	a. <i>iris</i> (iris) b. <i>bakul</i> (bakul)
6.	<i>peN-</i>	a. <i>pe-</i> b. <i>pem-</i> c. <i>pen-</i> d. <i>peng-</i> e. <i>peny-</i>	a. <i>penaseat</i> (penasihat) b. <i>pemasu</i> (pembasuh) c. <i>peniduk</i> (penidur) d. <i>pengatur</i> (pengatur) e. <i>penyukur</i> (penyukur)	a. <i>naseat</i> (nasihat) b. <i>basu</i> (basuh) c. <i>tiduk</i> (tidur) d. <i>atur</i> (atur) e. <i>cukur</i> (cukur)

Selain aturan prefiks tersebut, terdapat aturan kombinasi imbuhan kata yang mengandung lebih dari 1 awalan seperti pada tabel 4.3 berikut.

**Tabel 4. 3 Analisa Prefiks *se-* dan *peN-* Pada Konfiks**

No	Kata Imbuhan	Prefiks 1	Prefiks 2	Kata Dasar	Sufiks 1
1.	<i>sepenengeran</i> (sependengaran)	<i>se-</i>	<i>peN-</i>	<i>denger</i> (dengar)	<i>-an</i>

dalam contoh 1 tersebut terdapat 2 tingkat prefiks yaitu *se-* (prefiks 1) dan *peN-* (prefiks 2). Akan tetapi setelah dianalisa kembali, prefiks utama *peN-* merupakan gabungan awalan *pe-* dan *N-*. Sehingga dilakukan pemisahan prefiks dalam 3 tingkatan yaitu *se-* (prefiks 1), *pe-* (prefiks 2), dan *N-* (prefiks 3) seperti pada tabel 4.4 berikut.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 4. 4 Analisa Prefiks *se-*, *pe-*, dan *N-* Pada Konfiks**

No	Kata Imbuhan	Prefiks 1	Prefiks 2	Prefiks 3	Kata Dasar	Sufiks 1
1.	<i>sepenengaran</i> (sependengaran)	<i>se-</i>	<i>pe-</i>	<i>N-</i>	<i>denger</i> (dengar)	<i>-an</i>

Berdasarkan tabel 4.4 tersebut maka disusun aturan penghapusan prefiks khusus yang terdiri dari *pe-*, dan *se-* dimana 2 prefiks ini akan melakukan perulangan penghapusan prefiks 1 hingga 3 tingkat yang akan dijelaskan pada tahap perancangan. Selanjutnya, pada kata uji ditemukan kombinasi imbuhan yang tidak terdapat pada buku aturan tata bahasa untuk *Baso Pelembang* seperti pada tabel 4.5 berikut.

**Tabel 4. 5 Analisa Prefiks *me-*, *N-*, dan *ku-* Pada Konfiks**

No	Kata Imbuhan	Prefiks 1	Prefiks 2	Prefiks 3	Kata Dasar	Sufiks 1
1.	<i>mengajaknyo</i> (mengajaknya)	<i>me-</i>	<i>N-</i>	-	<i>ajak</i> (ajak)	<i>-nyo</i>
2.	<i>kujadike</i> (kujadikan)	<i>ku-</i>	-	-	<i>jadi</i> (jadi)	<i>-ke</i>

dalam tabel 4.5 nomor 1 tersebut terdapat 2 prefiks yaitu *me-* (prefiks 1) dan *N-* (prefiks 2). Berdasarkan tabel 4.5 nomor 1 tersebut maka disusun aturan penghapusan prefiks khusus yaitu *me-* dimana prefiks ini akan melakukan perulangan penghapusan prefiks 1 hingga 2 tingkat yang akan dijelaskan pada tahap perancangan. Selain itu, pada tabel 4.5 nomor 2 terdapat tambahan prefiks untuk aturan penghapusan prefiks yaitu *ku-*.

Beberapa simbol digunakan dalam aturan ini yaitu V untuk menjelaskan huruf setelah imbuhan berupa huruf vokal, C untuk menjelaskan huruf setelah imbuhan berupa huruf konsonan, dan A untuk menjelaskan huruf setelah imbuhan berupa huruf vokal atau konsonan. Aturan penghapusan prefiks pada tabel 4.6 sebagai berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 4. 6 Aturan Penghapusan Prefiks**

Aturan	Konstruksi	Penghapusan Prefiks
1.	{pe se} A...	{pe se} -A...
2.	{be ke te ng di ku m n} A...	{be ke te ng di ku m n} A...
3.	berV...	ber-V...
4.	ngeC...	nge-C...
5.	mV...	m- {b p} V...
6.	nV...	n- {d t} V...
7.	ngV...	ng- {g j k} V...
8.	nyV...	ny- {c j k s} V...
9.	meA...	me-A...
10.	penA...	pen-A...
11.	{b d t s} V...	{b d t s} -V...

#### b. Aturan Penghapusan Sufiks

Pada aturan penghapusan sufiks akan dilakukan analisa aturan penghapusan imbuhan yang terletak diakhir kata. Terdapat 6 sufiks utama yakni **-ke**, **-an**, **-i**, **-nyo**, **-la**, dan **-lah**. Sufiks ini memiliki beberapa wujud seperti yang dijabarkan pada tabel 4.7 berikut.

**Tabel 4. 7 Wujud Sufiks Menjadi Kata Dasar**

No	Kata Imbuhan	Kata Dasar	Sufiks
1.	<i>gantunke</i> (gantungkan)	<i>gantung</i> (gantung)	<b>-ke</b>
2.	<i>taunan</i> (tahunan)	<i>taun</i> (tahun)	<b>-an</b>
3.	<i>racuni</i> (racuni)	<i>racun</i> (racun)	<b>-i</b>
4.	<i>tingginyo</i> (tingginya)	<i>tinggi</i> (tinggi)	<b>-nyo</b>
5.	<i>sedila</i> (sedihlah)	<i>sedi</i> (sedih)	<b>-la</b>
6.	<i>pegilah</i> (pergilah)	<i>pegi</i> (pergi)	<b>-lah</b>

Setelah dilakukan pengumpulan kata uji, ditemukan 2 sufiks tambahan dalam aturan yaitu sufiks **-ku** dan **-mu** yang merupakan bentuk dari kepemilikan seperti pada tabel 4.8 berikut.

**Tabel 4. 8 Analisa Sufiks -mu dan -ku**

No	Kata Imbuhan	Kata Dasar	Sufiks
1.	<i>abahmu</i> (abahmu)	<i>abah</i> (abah)	<b>-mu</b>
2.	<i>anakku</i> (anakku)	<i>anak</i> (anak)	<b>-ku</b>

Selain aturan sufiks tersebut, terdapat aturan kombinasi imbuhan kata yang mengandung lebih dari 1 akhiran seperti pada tabel 4.9 berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 4. 9 Analisa Sufiks -ke, -nyo, dan -la Pada Konfiks**

No	Kata Imbuhan	Prefiks	Kata Dasar	Sufiks 1	Sufiks 2	Sufiks 3
1.	<i>diomongkenyola</i> (diomongkannyalah)	<i>di-</i>	<i>omong</i> (omong)	<i>-ke</i>	<i>-nyo</i>	<i>-la</i>

dalam tabel 4.9 tersebut terdapat 3 tingkat sufiks yaitu **-ke** (sufiks 1), **-nyo** (sufiks 2), dan **-la** (sufiks 3). Sehingga penghapusan sufiks akan dilakukan secara berulang 3 kali. Aturan penghapusan sufiks pada tabel 4.10 sebagai berikut.

**Tabel 4. 10 Aturan Penghapusan Sufiks**

Aturan	Konstruksi	Penghapusan Sufiks
12.	...A {ke an i ku mu nyo la lah}	...A-{ke an i ku mu nyo la lah}

c. Aturan Penghapusan Infiks

Pada aturan penghapusan infiks akan dilakukan analisa aturan penghapusan imbuhan yang terletak diawal kata setelah huruf pertama. Terdapat 3 infiks yakni **-el-**, **-em-**, dan **-er** seperti yang dijabarkan pada tabel 4.11 berikut.

**Tabel 4. 11 Wujud Infiks Menjadi Kata Dasar**

No	Kata Imbuhan	Infiks	Kata Dasar
1.	<i>telunjuk</i> (telunjuk)	<b>-el-</b>	<i>tunjuk</i> (tunjuk)
2.	<i>kemilau</i> (kemilau)	<b>-em-</b>	<i>kilau</i> (kilau)
3.	<i>geretak</i> (geretak)	<b>-er-</b>	<i>getak</i> (getak)

Aturan penghapusan infiks pada tabel 4.12 sebagai berikut.

**Tabel 4. 12 Aturan Penghapusan Infiks**

Aturan	Konstruksi	Penghapusan Infiks
13.	C {el} V...	C-{el}-V...
14.	C {em} V...	C-{em}-V...
15.	C {er} V...	C-{er}-V...

### 4.1.3 Analisa Algoritma Stemming Baso Pelembang

Pada tahapan analisa algoritma *stemming Baso Pelembang* diketahui proses penghapusan imbuhan (afiks) terdiri dari imbuhan awal (prefiks), imbuhan akhir



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(sufiks), dan imbuhan sisipan (infiks), serta kombinasi imbuhan (konfiks) yang merupakan gabungan dari prefiks dan sufiks. Berikut ini merupakan penjelasan proses penghapusan afiks untuk algoritma *stemming Baso Pelembang*:

#### 1. Penghapusan Imbuhan Awal (Prefiks)

Proses penghapusan imbuhan awal (prefiks) dilakukan sesuai aturan penghapusan yang telah disusun seperti pada tabel 4.13 berikut ini.

**Tabel 4. 13 Proses Penghapusan Prefiks**

Proses	Kata	Prefiks 1
1	<i>maco</i> (membaca)	<i>N-</i>
2	<i>baco</i> (baca)	-

Pada kata *maco* dilakukan penghapusan prefiks sebanyak 1 kali karena kata memiliki 1 tingkatan prefiks yaitu *N-*. Proses 1 kata *maco* akan memasuki tahap pengecekan kata pada *database* untuk mengetahui apakah kata tersedia atau tidak pada kamus. Hasilnya, kata tidak tersedia pada kamus maka akan dilakukan proses penghapusan imbuhan. Tahapan awal algoritma akan menghapus prefiks paling awal yang terletak pada urutan prefiks 1 sehingga kata *maco* menjadi kata *baco*. Proses 2 menunjukkan kata *baco* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata *baco* merupakan kata dasar yang tersedia pada kamus sehingga proses penghapusan imbuhan berhenti (selesai).

#### 2. Penghapusan Imbuhan Akhir (Sufiks)

Proses penghapusan imbuhan akhir (sufiks) dilakukan sesuai aturan penghapusan yang telah disusun seperti pada tabel 4.14 berikut ini.

**Tabel 4. 14 Proses Penghapusan Sufiks**

Proses	Kata	Sufiks 1
1	<i>abahmu</i> (ayahmu)	<i>-mu</i>
2	<i>aba</i> (ayah)	-

Pada kata *abahmu* dilakukan penghapusan sufiks sebanyak 1 kali karena kata memiliki 1 tingkatan sufiks yaitu *-mu*. Proses 1 kata *abahmu* akan

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

memasuki tahap pengecekan kata pada *database* untuk mengetahui apakah kata tersedia atau tidak pada kamus. Hasilnya, kata tidak tersedia pada kamus maka akan dilakukan proses penghapusan imbuhan. Tahapan awal algoritma akan menghapus sufiks paling awal yang terletak pada urutan sufiks 1 sehingga kata *abahmu* menjadi kata *aba*. Proses 2 menunjukkan kata *aba* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata *aba* merupakan kata dasar yang tersedia pada kamus sehingga proses penghapusan imbuhan berhenti (selesai).

#### 3. Penghapusan Imbuhan Sisipan (Infiks)

Proses penghapusan imbuhan sisipan (infiks) dilakukan sesuai aturan penghapusan yang telah disusun. Terdapat 1 tingkatan proses penghapusan infiks terpanjang yang ditemukan seperti pada tabel 4.15 berikut ini.

**Tabel 4. 15 Proses Penghapusan Infiks**

Proses	Kata	Infiks 1
1	<i>telunjuk</i> (telunjuk)	<b>-el-</b>
2	<i>tunjuk</i> (tunjuk)	-

Pada kata *telunjuk* dilakukan penghapusan infiks sebanyak 1 kali karena kata memiliki 1 tingkatan infiks yaitu **-el-**. Proses 1 kata *telunjuk* akan memasuki tahap pengecekan kata pada *database* untuk mengetahui apakah kata tersedia atau tidak pada kamus. Hasilnya, kata tidak tersedia pada kamus maka akan dilakukan proses penghapusan imbuhan. Tahapan awal algoritma akan menghapus infiks paling awal yang terletak pada urutan infiks 1 yaitu setelah huruf *t* huruf pertama pada kata *telunjuk* terdapat wujud imbuhan **-el-** sehingga kata *telunjuk* menjadi *tunjuk*. Proses 2 menunjukkan kata *tunjuk* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata *tunjuk* merupakan kata dasar yang tersedia pada kamus sehingga proses penghapusan imbuhan berhenti(selesai).

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

#### 4. Penghapusan Kombinasi Imbuhan (Konfiks)

Proses penghapusan kombinasi imbuhan (konfiks) dilakukan sesuai aturan penghapusan prefiks dan sufiks yang dilakukan secara bertahap dari penghapusan prefiks lalu sufiks seperti pada tabel 4.16 berikut ini.

**Tabel 4. 16 Proses Penghapusan Konfiks**

Proses	Kata	Prefiks 1	Sufiks 1	Sufiks 2	Sufiks 3
1	<i>diomongkenyola</i> (diomongkannyalah)	<i>di-</i>	<i>-ke</i>	<i>-nyo</i>	<i>-la</i>
2	<i>omongkenyola</i> (omongkannyalah)	-	<i>-ke</i>	<i>-nyo</i>	<i>-la</i>
3	<i>omongkenyo</i> (omongkannya)	-	<i>-ke</i>	<i>-nyo</i>	-
4	<i>Omongke</i> (Omongkan)	-	<i>-ke</i>	-	-
5	<i>Omong</i> (Omong)	-	-	-	-

Pada kata *diomongkenyola* dilakukan penghapusan prefiks dilakukan sebanyak 1 kali karena kata memiliki 1 tingkatan prefiks yaitu *di-* serta penghapusan sufiks sebanyak 3 kali karena kata memiliki 3 tingkatan sufiks yaitu *-ke*, *-nyo*, dan *-la*. Proses 1 kata *diomongkenyola* akan memasuki tahap pengecekan kata pada *database* untuk mengetahui apakah kata tersedia atau tidak pada kamus. Hasilnya, kata tidak tersedia pada kamus maka akan dilakukan proses penghapusan imbuhan. Tahapan awal algoritma akan menghapus prefiks yang terletak pada urutan prefiks 1 yaitu *di-* sehingga kata *diomongkenyola* menjadi *omongkenyola*. Proses 2 kata *omongkenyola* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata tidak tersedia pada kamus maka akan dilakukan proses penghapusan imbuhan. Algoritma akan menghapus sufiks paling akhir yang terletak pada urutan sufiks 3 yaitu *-la* sehingga kata *omongkenyola* menjadi *omongkenyo*. Proses 3 kata *omongkenyo* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata tidak tersedia pada kamus maka dilakukan proses penghapusan imbuhan. Algoritma akan menghapus sufiks paling akhir yang terletak pada urutan sufiks 2 yaitu *-nyo* sehingga kata *omongkenyo* menjadi *omongke*. Proses 4 kata *omongke* akan memasuki tahap pengecekan kata pada *database*.





#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hasilnya, kata tidak tersedia pada kamus maka dilakukan proses penghapusan imbuhan. Algoritma akan menghapus sufiks paling akhir yang terletak pada urutan sufiks 1 yaitu *-ke* sehingga kata *omongke* menjadi *omong*. Proses 5 menunjukkan kata *omong* akan memasuki tahap pengecekan kata pada *database*. Hasilnya, kata *omong* merupakan kata dasar yang tersedia pada kamus sehingga proses penghapusan imbuhan berhenti (selesai).

## 4.2 Perancangan Algoritma

Pada tahap perancangan algoritma akan dijelaskan perancangan *flowchart* algoritma *stemming* dan perancangan fungsi cek kamus, fungsi cek prefiks, fungsi hapus prefiks, fungsi cek sufiks, fungsi hapus sufiks, fungsi cek infiks, fungsi hapus infiks, serta fungsi *stemming* dalam bentuk *pseudocode* sebagai berikut.

### 4.2.1 Perancangan Flowchart Algoritma Stemming

Perancangan *flowchart* algoritma *stemming* dibuat sebagai gambaran alur kerja dari algoritma *stemming* untuk *Baso Pelembang*. *Flowchart* ini dibuat berdasarkan hasil analisa algoritma.

*Flowchart* dimulai dari memasukkan kata dalam *Baso Pelembang*. Kata akan diproses pada tahap pengecekan awal yaitu tahap pengecekan kata pada kamus yang terletak di *database*. Jika kata ditemukan pada kamus, maka kata merupakan kata dasar. Proses *stemming* tidak dilakukan dan tahapan selesai. Jika kata tidak ditemukan pada kamus, maka proses *stemming* akan dilakukan dengan urutan tahap penghapusan dimulai dari prefiks, sufiks, lalu infiks. Apabila kata yang telah hapus imbuhan terdapat pada kamus, maka kata merupakan kata dasar. Namun, jika kata yang telah dihapus imbuhan tidak terdapat pada kamus, maka kata akan dikembalikan menjadi kata original seperti pada *flowchart* berikut ini.

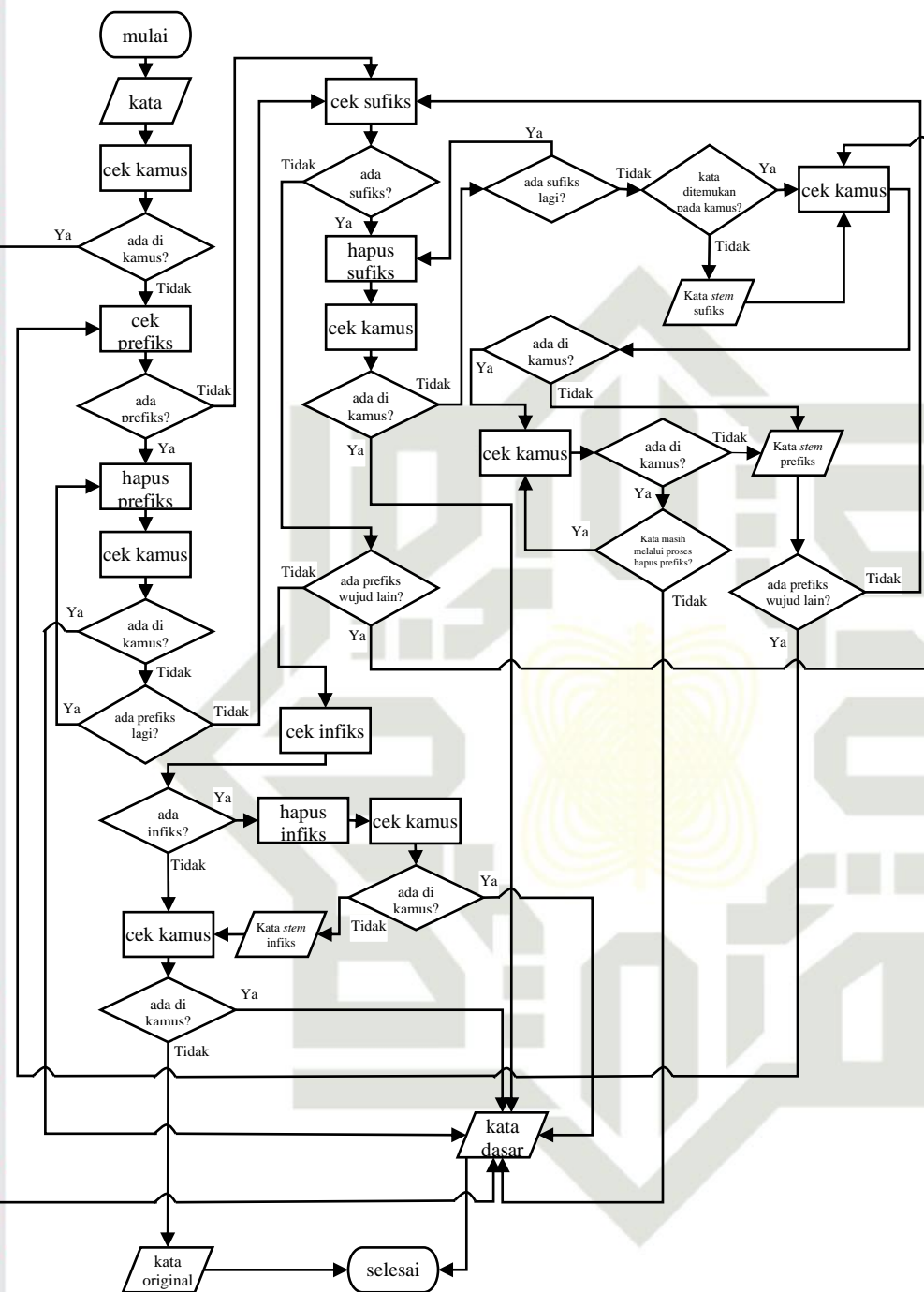
UIN SUSKA RIAU

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4. 5 Flowchart Algoritma Stemming Baso Pelembang

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Penjelasan tahapan *flowchart*:

1. Proses dimulai ketika kata dimasukkan pada sistem.
2. Kata yang telah dimasukkan dicek pada kamus yang terdapat dalam *database*.
3. Jika kata terdapat pada kamus, maka kata dinyatakan sebagai kata dasar dan proses selesai.
4. Jika kata tidak ditemukan pada kamus, maka dilakukan cek prefiks.
5. Jika kata terdapat prefiks, maka akan dilakukan penghapusan prefiks.
6. Jika kata tidak terdapat prefiks, maka akan dilakukan cek sufiks.
7. Setelah prefiks dihapus, kata akan dicek pada kamus.
8. Jika kata ditemukan pada kamus, maka kata merupakan kata dasar dan tahapan selesai.
9. Jika kata tidak ditemukan pada kamus, maka akan dilakukan cek prefiks kembali untuk memastikan kata berupa prefiks atau konfiks.
10. Jika kata masih terdapat prefiks, maka dilakukan penghapusan prefiks kembali.
11. Jika kata tidak terdapat prefiks, maka akan dilakukan proses cek sufiks.
12. Tahapan selanjutnya adalah cek sufiks.
13. Jika kata terdapat sufiks, maka akan dilakukan penghapusan sufiks.
14. Jika kata tidak terdapat sufiks, maka akan dilakukan cek prefiks dalam wujud lain (contohnya prefiks dengan awal *te-* yang memiliki wujud *te-* dan *t-*).
15. Setelah sufiks dihapus, kata akan dicek pada kamus.
16. Jika kata ditemukan pada kamus, maka kata merupakan kata dasar dan tahapan selesai.
17. Jika kata tidak ditemukan pada kamus, maka akan dilakukan cek sufiks kembali untuk memastikan kata berupa sufiks atau konfiks.
18. Jika kata masih terdapat sufiks, maka dilakukan penghapusan sufiks kembali.
19. Jika kata tidak terdapat sufiks namun merupakan kata konfiks, maka akan dilakukan cek ketersediaan kata pada kamus.
20. Jika kata terdapat pada kamus, maka dalam proses hapus sufiks dilakukan cek kata kembali pada kamus untuk tahap setelah hapus sufiks.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

21. Jika kata tidak terdapat pada kamus, maka sufiks dikembalikan pada kata yang disebut kata *stem* sufiks.
22. Setelah sufiks dikembalikan, dilakukan cek ketersediaan kata pada kamus.
23. Jika kata terdapat pada kamus, maka dalam proses hapus sufiks dilakukan cek kata kembali pada kamus untuk tahap setelah hapus sufiks.
24. Cek kata pada kamus untuk tahap hapus sufiks.
25. Jika kata terdapat pada kamus, maka dilanjutkan cek kata pada kamus untuk tahap hapus prefiks.
26. Jika kata tidak terdapat pada kamus, maka prefiks dikembalikan.
27. Cek kata pada kamus untuk tahap hapus prefiks.
28. Jika terdapat kata pada kamus, dilakukan cek kata pada kamus kembali untuk prefiks (jika prefiks lebih dari 1 pada konfiks).
29. Jika kata tidak terdapat pada kamus, maka prefiks dikembalikan pada kata yang disebut kata *stem* prefiks.
30. Jika kata sudah tidak melalui cek kamus pada proses hapus prefiks dan kata terdapat pada kamus, maka proses berhenti dan kata merupakan kata dasar.
31. Jika prefiks dikembalikan, selanjutnya dilakukan cek wujud prefiks yang lain. Jika ada prefiks dalam wujud lain, maka tahapan cek prefiks dilakukan kembali.
32. Jika tidak ada prefiks dalam wujud lain, maka dilanjutkan pada tahap cek sufiks.
33. Pada tahap cek sufiks jika tidak ada prefiks dalam wujud lain, maka dilakukan cek infiks.
34. Pada tahapan cek infiks jika kata terdapat infiks, maka akan dilakukan penghapusan infiks.
35. Jika kata tidak terdapat infiks, maka akan dilakukan cek kata pada kamus.
36. Setelah infiks dihapus, kata akan dicek pada kamus.
37. Jika kata ditemukan pada kamus, maka kata merupakan kata dasar dan tahapan selesai.
38. Jika kata tidak ditemukan pada kamus, maka infiks dikembalikan pada kata yang disebut kata *stem* infiks.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

39. Setelah infiks dikembalikan, dilakukan cek kata pada kamus.

40. Jika kata terdapat pada kamus, maka kata merupakan kata dasar dan tahapan selesai.

41. Jika kata tidak terdapat pada kamus, maka kata merupakan kata original yaitu kata dikembalikan menjadi kata pertama saat dimasukkan (kata asal) dan tahapan selesai.

#### 4.2.2 Perancangan Fungsi Cek Kamus

Fungsi cek kamus dirancang untuk melakukan pengecekan kata yang dimasukkan pada sistem merupakan kata dasar yang tersedia pada kamus atau tidak. Tahap cek kamus ini dilakukan sebelum dan selama proses *stemming* seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi cek kamus ditampilkan dengan nama fungsi *cekKamus()* dalam bentuk *pseudocode* sebagai berikut.

```
function cekKamus(kata)
{ cari di database }
  if kata ada pada kamus then
    return true { true jika ada }
  else
    return fase { false jika tidak ada }
  endif
endfunction
```

**Gambar 4. 6 Pseudocode Fungsi Cek Kamus**

Pada perancangan *pseudocode* untuk cek kamus jika kata yang dicari ditemukan dalam kamus, maka proses pencarian imbuhan akan berhenti. Jika kata belum ditemukan maka kata diproses pada tahap *stemming* yang terdiri dari fungsi hapus prefiks, fungsi hapus sufiks, dan fungsi hapus infiks.

#### 4.2.3 Perancangan Fungsi Cek Prefiks

Fungsi cek prefiks dirancang untuk melakukan pengecekan kata yang dimasukkan pada sistem merupakan kata yang memiliki imbuhan awal atau tidak. Tahap cek prefiks ini dilakukan dalam proses *stemming* seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi cek prefiks ditampilkan dengan nama fungsi *cekPrefiks()* dalam bentuk *pseudocode* sebagai berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```
function cekPrefiks(kata)
{ cari pola kata }
  if kata memiliki awalan then
    kata <- hapusPrefiks(kata)
  else
    kata <- cekSufiks(kata)
  endif
endfunction
```

**Gambar 4. 7 Pseudocode Fungsi Cek Prefiks**

Pada perancangan *pseudocode* untuk cek prefiks jika kata yang dicari ditemukan imbuhan awal, maka proses akan dilanjutkan dengan fungsi *hapusPrefiks()*. Jika kata tidak ditemukan imbuhan awal, maka kata memasuki fungsi *cekSufiks()*.

#### 4.2.4 Perancangan Fungsi Hapus Prefiks

Fungsi hapus prefiks dirancang untuk melakukan penghapusan imbuhan awal pada kata. Tahap hapus prefiks ini dilakukan selama proses *stemming* berlangsung dan selama prefiks masih terdapat pada kata seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi hapus prefiks ditampilkan dengan nama fungsi *hapusPrefiks()* dalam bentuk *pseudocode* sebagai berikut.

```
function hapusPrefiks(kata)

  if cari pola awalan pada kata then
    kata <- hapus awalan pada kata
    if cekKamus(kata)
      return kata
    endif
  endif
endfunction
```

**Gambar 4. 8 Pseudocode Fungsi Hapus Prefiks**

Pada perancangan *pseudocode* fungsi hapus prefiks, terdapat 11 pola awalan yang merupakan bagian dari 11 aturan penghapusan prefiks yang telah dijelaskan pada tabel 4.3 aturan penghapusan prefiks. Jika pola awalan ditemukan, maka awalan pada kata akan dihapus. Setelah prefiks dihapus, dilakukan pengecekan kata pada kamus.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 4.2.5 Perancangan Fungsi Cek Sufiks

Fungsi cek sufiks dirancang untuk melakukan pengecekan kata yang dimasukkan pada sistem merupakan kata yang memiliki imbuhan akhir atau tidak. Tahap cek sufiks ini dilakukan dalam proses *stemming* seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi cek sufiks ditampilkan dengan nama fungsi *cekSufiks()* dalam bentuk *pseudocode* sebagai berikut.

```
function cekSufiks(kata)
{ cari pola kata }
  if kata memiliki akhiran then
    kata <- hapusSufiks(kata)
  elseif
    kata <- cekPrefiks(kata)
  else
    kata <- cekInfiks(kata)
  endif
endfunction
```

**Gambar 4. 9 Pseudocode Fungsi Cek Sufiks**

Pada perancangan *pseudocode* untuk cek sufiks jika kata yang dicari ditemukan imbuhan akhir, maka proses akan dilanjutkan dengan fungsi *hapusSufiks()*. Jika kata tidak ditemukan imbuhan akhir, maka kata memasuki fungsi *cekPrefiks()* untuk memastikan tidak ada prefiks dalam wujud lain. Jika tidak ada prefiks dalam wujud lain, maka kata memasuki fungsi *cekInfiks()*.

### 4.2.6 Perancangan Fungsi Hapus Sufiks

Fungsi hapus sufiks dirancang untuk melakukan penghapusan imbuhan akhir pada kata. Tahap hapus sufiks ini dilakukan selama proses *stemming* berlangsung dan selama sufiks masih terdapat pada kata seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi hapus sufiks ditampilkan dengan nama fungsi *hapusSufiks()* dalam bentuk *pseudocode* sebagai berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```
function hapusSufiks(kata)

    if cari pola akhiran pada kata then
        kata <- hapus akhiran pada kata
        if cekKamus(kata)
            return kata
        endif
    endif
endfunction
```

**Gambar 4. 10 Pseudocode Fungsi Hapus Sufiks**

Pada perancangan *pseudocode* fungsi hapus sufiks, terdapat 1 pola akhiran yang merupakan bagian dari aturan penghapusan sufiks yang telah dijelaskan pada tabel 4.4 aturan penghapusan sufiks. Jika pola akhiran ditemukan, maka akhiran pada kata akan dihapus. Setelah sufiks dihapus, dilakukan pengecekan kata pada kamus.

#### 4.2.7 Perancangan Fungsi Cek Infiks

Fungsi cek infiks dirancang untuk melakukan pengecekan kata yang dimasukkan pada sistem merupakan kata yang memiliki imbuhan sisipan atau tidak. Tahap cek infiks ini dilakukan dalam proses *stemming* seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi cek infiks ditampilkan dengan nama fungsi *cekInfiks()* dalam bentuk *pseudocode* sebagai berikut.

```
function cekInfiks(kata)

{ cari pola kata }
    if kata memiliki sisipan then
        kata <- hapusInfiks(kata)
    else
        kata <- cekKamus(kata)
    endif
endfunction
```

**Gambar 4. 11 Pseudocode Fungsi Cek Infiks**

Pada perancangan *pseudocode* untuk cek infiks jika kata yang dicari ditemukan imbuhan sisipan, maka proses akan dilanjutkan dengan fungsi *hapusInfiks()*. Jika kata tidak ditemukan imbuhan sisipan, maka kata memasuki fungsi *cekKamus()* untuk memastikan ketersediaan hasil kata pada kamus.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 4.2.8 Perancangan Fungsi Hapus Infiks

Fungsi hapus infiks dirancang untuk melakukan penghapusan imbuhan sisipan pada kata. Tahap hapus infiks ini dilakukan selama proses *stemming* berlangsung seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi hapus infiks ditampilkan dengan nama fungsi *hapusInfiks()* dalam bentuk *pseudocode* sebagai berikut.

```
function hapusInfiks(kata)
    if cari pola sisipan pada kata then
        kata <- hapus sisipan pada kata
        if cekKamus(kata)
            return kata
        endif
    endif
endfunction
```

**Gambar 4. 12 Pseudocode Fungsi Hapus Infiks**

Pada perancangan *pseudocode* fungsi hapus infiks, terdapat 1 pola sisipan yang merupakan bagian dari aturan penghapusan infiks yang telah dijelaskan pada tabel 4.5 aturan penghapusan infiks. Jika pola sisipan ditemukan, maka sisipan pada kata akan dihapus. Setelah infiks dihapus, dilakukan pengecekan kata pada kamus.

### 4.2.9 Perancangan Fungsi *Stemming*

Fungsi hapus *stemming* dirancang untuk menjalankan seluruh fungsi penghapusan imbuhan secara bertahap. Tahap fungsi *stemming* ini dilakukan sejak kata dimasukkan pada sistem dan melewati proses *stemming* seperti yang telah dijelaskan pada *flowchart* algoritma *stemming Baso Pelembang*. Perancangan fungsi *stemming* ditampilkan dengan nama fungsi *stemming()* dalam bentuk *pseudocode* sebagai berikut.



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```
function stemming(kata)
    if cekKamus(kata) then
        return kata
    else
        if cekPrefiks(kata) then
            kata <- hapusPrefiks(kata)
        else
            if cekSufiks(kata) then
                kata <- hapusSufiks(kata)
            else
                if cekInfiks(kata) then
                    kata <- hapusInfiks(kata)
                else
                    kata <- cekKamus(kata)
                    return kata
                endif
            endif
        endif
    endif
endfunction
```

**Gambar 4. 13 Pseudocode Fungsi Stemming**

Pada perancangan *pseudocode* fungsi *stemming*, fungsi yang dijalankan pertama adalah fungsi *cekKamus()* untuk mengecek ketersediaan kata pada kamus dalam *database*. Jika kata tersedia maka proses berhenti dan kata dinyatakan sebagai kata dasar. Jika kata tidak tersedia pada kamus, dijalankan fungsi *cekPrefiks()*. Jika kata memiliki prefiks, maka dijalankan fungsi *hapusPrefiks()*. Jika tidak memiliki prefiks, maka dijalankan fungsi *cekSufiks()*. Jika kata memiliki sufiks, maka dijalankan fungsi *hapusSufiks()*. Jika tidak memiliki sufiks, maka dijalankan fungsi *cekInfiks()*. Jika kata memiliki infiks, maka dijalankan fungsi *hapusInfiks()*. Jika kata tidak memiliki infiks, maka dijalankan fungsi *cekKamus()* kembali. Jika kata tidak ditemukan pada kamus, maka proses berhenti dan kata tersebut dikembalikan menjadi kata awal dimasukkan. Kata tersebut kemudian disebut kata original.

### 4.3 Perancangan Database

Pada tahap perancangan *database* untuk menyimpan data kamus yang berisi kata dasar dan data dokumen berbahasa daerah yang berisi kata uji maka dibuat *database* dengan nama *pengumpulan\_data*. Pada *database* *pengumpulan\_data* dirancang 2 struktur tabel yang terdiri dari tabel *kata\_dasar* dan tabel *kata\_uji* yang dijelaskan sebagai berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Tabel 4. 17 Perancangan Struktur Tabel *kata\_dasar***

Nama Field	Type	Null	Default	Extra
Dasar_id	int(11)	No	None	AUTO_INCREMENT
Kata	tinytext	No	None	
Arti	tinytext	No	None	

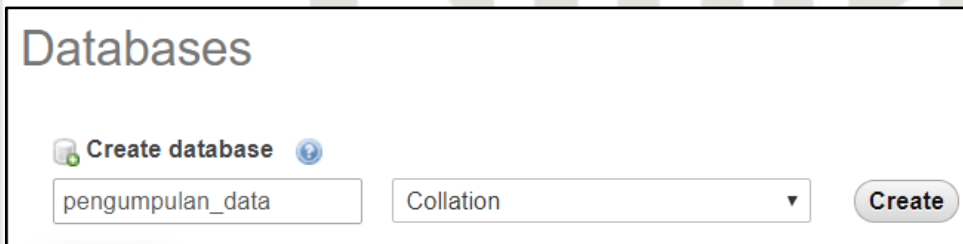
Pada perancangan struktur tabel *kata\_dasar* dibuat 3 *field* yang terdiri dari *field Dasar\_id* sebagai *Primary Key*, *field Kata* yang berisi kumpulan kata dalam *Baso Pelembang*, dan *field Arti* yang berisi kumpulan arti dari kata dalam *Baso Pelembang*.

**Tabel 4. 18 Perancangan Struktur Tabel *kata\_uji***

Nama Field	Type	Null	Default	Extra
Uji_id	int(11)	No	None	AUTO_INCREMENT
Kata_Uji	tinytext	No	None	

Pada perancangan struktur tabel *kata\_uji* dibuat 2 *field* yang terdiri dari *field Uji\_id* sebagai *primary key* dan *field Kata\_Uji* yang berisi kumpulan kata uji dalam *Baso Pelembang*.

Setelah struktur tabel dirancang, akan dilakukan pemindahan data yang telah dikumpulkan pada *ms.excel* ke *database pengumpulan\_data* dengan langkah sebagai berikut.



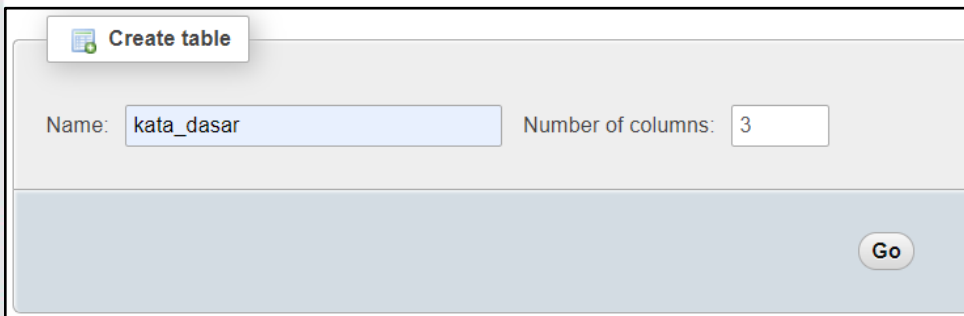
The screenshot shows the 'Databases' section of a MySQL management tool. There is a 'Create database' button with a plus icon. Below it, a text input field contains 'pengumpulan\_data'. To the right of the input field is a 'Collation' dropdown menu. Further right is a 'Create' button.

**Gambar 4. 14 Perancangan Database**

Tahap pertama untuk perancangan *database* yaitu akan dilakukan pembuatan sebuah *database* dengan nama *pengumpulan\_data*. Perancangan *database* ini menggunakan aplikasi XAMPP sebagai server dengan cara mengakses *localhost/phpmyadmin*. Setelah nama *database* diisi, pilih *create*. Maka *database pengumpulan\_data* siap untuk digunakan.

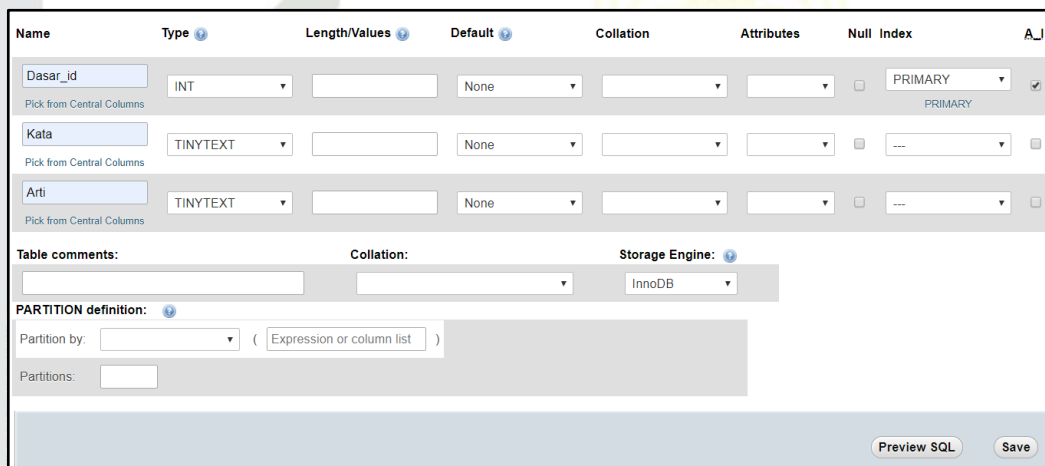
#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Gambar 4. 15 Perancangan Tabel Kata Dasar**

Tahap selanjutnya untuk perancangan *database* yaitu akan dilakukan pembuatan tabel untuk kata dasar dengan nama *kata\_dasar*. Pembuatan tabel ini masih menggunakan aplikasi XAMPP sebagai server dengan cara mengakses *localhost/phpmyadmin* kemudian pilih *database pengumpulan\_data*. Isi nama tabel dan *number of columns* sesuai rancangan yang telah dibuat. pilih *Go*, maka akan diarahkan pada halaman pembuatan *field*.



**Gambar 4. 16 Field untuk Tabel Kata Dasar**

Tahap selanjutnya pada halaman pembuatan *field* untuk tabel kata dasar yaitu mengisi setiap *field* sesuai perancangan yang telah dibuat. *Field Dasar\_id* dijadikan *primary key*. Setelah semua diisi, maka pilih *Save* dan tabel kata dasar siap untuk digunakan.

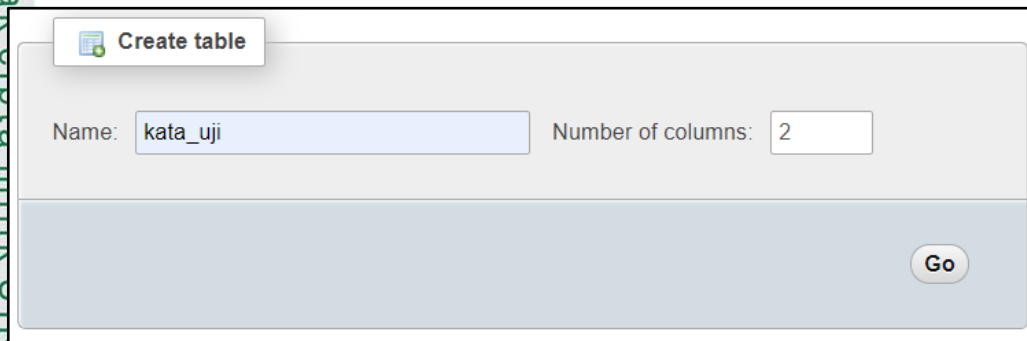


#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

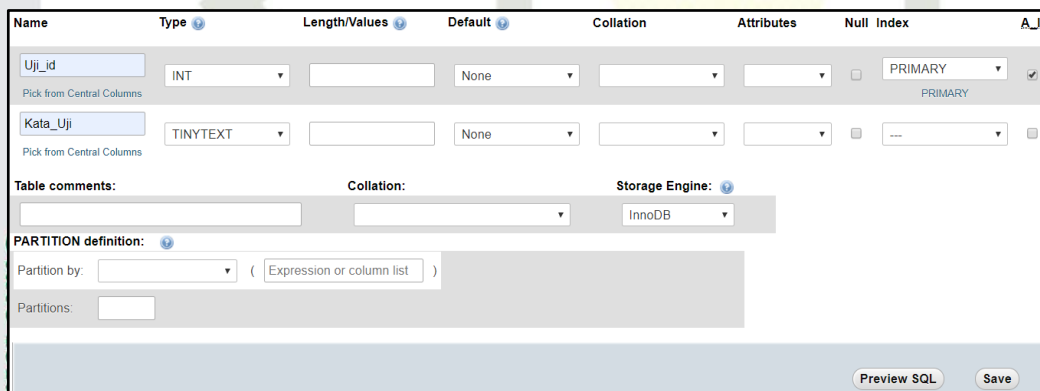
- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Gambar 4. 17 Perancangan Tabel Kata Uji**

Tahap selanjutnya untuk perancangan *database* yaitu akan dilakukan pembuatan tabel untuk kata uji dengan nama *kata\_uji* seperti pada tahapan pembuatan tabel kata dasar. Pembuatan tabel ini masih menggunakan aplikasi XAMPP sebagai server dengan cara mengakses *localhost/phpmyadmin* kemudian pilih *database pengumpulan\_data*. Isi nama tabel dan *number of columns* sesuai rancangan yang telah dibuat. Pilih *Go*, maka akan diarahkan pada halaman pembuatan *field*.



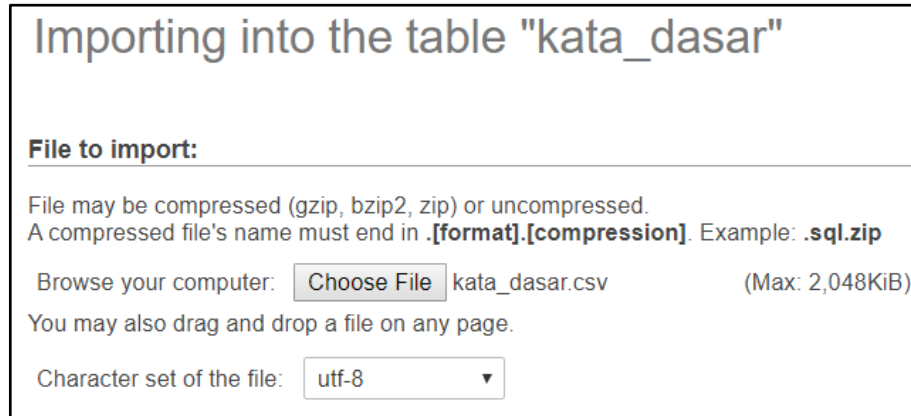
**Gambar 4. 18 Field untuk Tabel Kata Uji**

Tahap selanjutnya pada halaman pembuatan *field* untuk tabel kata uji yaitu mengisi setiap *field* sesuai perancangan yang telah dibuat. *Field Uji\_id* dijadikan *primary key*. Setelah semua diisi, maka pilih *Save* dan tabel kata uji siap untuk digunakan.

Pembuatan tabel kata dasar dan tabel kata uji telah selesai, namun tabel belum memiliki isi yang berupa kumpulan data. Kumpulan data yang telah dibuat pada *ms.excel* akan dimasukkan ke *database pengumpulan\_data* pada masing-masing tabel yang telah tersedia dengan cara sebagai berikut.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Importing into the table "kata\_dasar"

**File to import:**

File may be compressed (gzip, bzip2, zip) or uncompressed.  
A compressed file's name must end in **[format].[compression]**. Example: **.sql.zip**

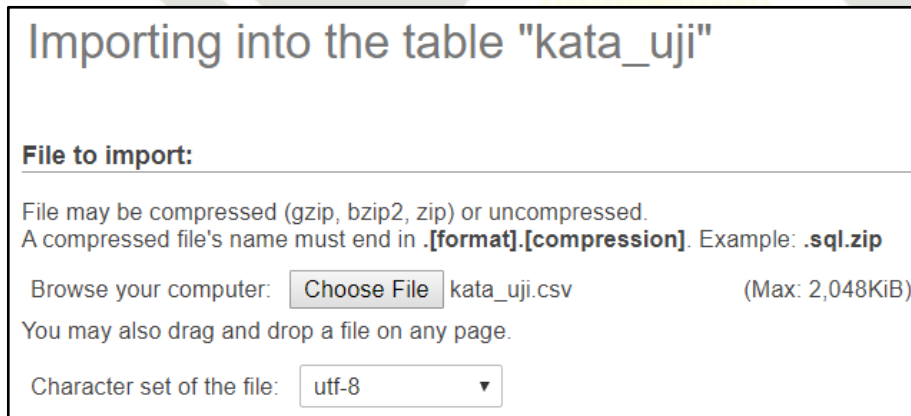
Browse your computer:  kata\_dasar.csv (Max: 2,048KiB)

You may also drag and drop a file on any page.

Character set of the file:

**Gambar 4. 19 Memasukkan Data Kata Dasar pada Database**

Pada tahap memasukkan data kata dasar pada *database* dilakukan dengan cara mengakses *localhost/phpmyadmin* pilih *database pengumpulan\_data*. Setelah itu, pilih tabel *kata\_dasar* dan klik *import*. Pada halaman *import*, tekan tombol *choose file*. Pilih file *kata\_dasar.csv* yang telah disediakan. *Scroll* halaman *import* sampai bawah dan pilih *Go*. Maka seluruh data kata dasar akan masuk pada *database pengumpulan\_data* dalam tabel *kata\_dasar*.



Importing into the table "kata\_uji"

**File to import:**

File may be compressed (gzip, bzip2, zip) or uncompressed.  
A compressed file's name must end in **[format].[compression]**. Example: **.sql.zip**

Browse your computer:  kata\_uji.csv (Max: 2,048KiB)

You may also drag and drop a file on any page.

Character set of the file:

**Gambar 4. 20 Memasukkan Data Kata Uji pada Database**

Pada tahap memasukkan data uji, sama seperti memasukkan data kata dasar. Pada *database* dilakukan dengan cara mengakses *localhost/phpmyadmin* pilih *database pengumpulan\_data*. Setelah itu, pilih tabel *kata\_uji* dan klik *import*. Pada halaman *import*, tekan tombol *choose file*. Pilih file *kata\_uji.csv* yang telah disediakan. *Scroll* halaman *import* sampai bawah dan pilih *Go*. Maka seluruh data kata uji akan masuk pada *database pengumpulan\_data* dalam tabel *kata\_uji*.



#### Hak Cipta Dilindungi Undang-Undang

## BAB VI KESIMPULAN

### 6.1 Kesimpulan

Berdasarkan hasil penelitian tentang algoritma *stemming* untuk *Baso Pelembang* berbasis aturan tata bahasa dapat disimpulkan hasil pengujian terhadap kata uji berhasil dijalankan sesuai perancangan dan hampir seluruh keluaran dari sistem sesuai dengan harapan yang diinginkan. Pengujian akurasi terhadap 645 kata berdasarkan aturan tata bahasa dan ketersediaan kata dasar pada kamus didapatkan sebesar 99.84% dengan 3 model kombinasi algoritma yang dapat dijalankan yaitu kombinasi 1, 3, dan 5. Akan tetapi, ditemukan 12 kata yang tidak sesuai harapan dan menyisakan 633 kata yang sesuai harapan berdasarkan hasil validasi oleh validator. Sehingga hasil akurasi akhir adalah sebesar 98.14%. Algoritma termasuk ke dalam kategori *heavy stemming*. Kegagalan dan kesalahan pada proses *stemming* (*overstemming*) terjadi karena algoritma masih memiliki beberapa ketidaksempurnaan dan ketersediaan kata pada kamus yang masih belum lengkap. Terdapat perbedaan akurasi sekitar 1.7% yang membuktikan bahwa algoritma dapat berjalan dengan baik.

### 6.2 Saran

Berdasarkan hasil kesimpulan dalam penelitian, terdapat saran yang dapat digunakan untuk penelitian selanjutnya yang lebih baik sebagai berikut:

1. Menurut validator, kamus yang digunakan belum lengkap. Sehingga perlu dilakukan penambahan kata pada kamus untuk meningkatkan hasil akurasi algoritma *stemming*.
2. Algoritma *stemming* ini dapat dikembangkan dan diimplementasikan menjadi aplikasi terjemahan kata pada dokumen dalam *Baso Pelembang* ke Bahasa Indonesia atau sebaliknya.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR PUSTAKA

- Adriani, M., Asian, J., Nazief, B., & Williams, H. E. (2007). Stemming Indonesian : A Confix Stripping Approach. *ACM Transactions on Asian Language Information Processing (TALIP)*, 6(4), 1–33. <https://doi.org/10.1145/1316457.1316459>.
- Agastya, I. M. A. (2018). Pengaruh Stemmer Bahasa Indonesia Terhadap Performa Analisis Sentimen Terjemahan Ulasan Film. *ResearchGate*, (February).
- Agusta, L. (2009). Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Andriani Untuk Stemming Dokumen Teks Bahasa Indonesia. *Konferensi Nasional Sistem dan Informatika*, 196–201. <https://doi.org/10.1016/j.ijdevneu.2011.05.008>
- Al Fatta, H. (2007). *Analisis dan Perancangan Sistem Informasi untuk Keunggulan Bersaing Perusahaan dan Organisasi Modern*. Yogyakarta: Penerbit ANDI.
- Aliana, S. A., Nursato, S., Arifin, S. S., Soetopo, S., & Waif, M. (1987). *Morfologi dan Sintaksis Bahasa Melayu Palembang*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan.
- Amin, B. M., Amin, A. A., Amin, M., & Tadjuddin, Z. A. M. (2009). *Tata Bahasa dan Kamus Baso Pelembang*. Palembang: Yayasan Madrasah Najahiyah.
- Barakbah, A. R., Karlita, T., & Ahsan, A. S. (2013). *Logika dan algoritma*. Surabaya.
- Cormen, T. H. (1989). *Introduction to Algorithms*. London: The MIT Press.
- Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1990). *Introduction to algorithms*. New York: McGraw-Hill Book Company.
- Hanafiah, D. (1995). *Melayu-Jawa: Citra Budaya dan Sejarah Palembang*. Jakarta: PT. Raja Grafindo Persada.
- Hapsari, R. K., & Santoso, Y. J. (2015). Stemming Artikel Berbahasa Indonesia dengan Pendekatan Confix-Stripping. *Prosiding Seminar Nasional Manajemen Teknologi XXII*, 1–8.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- Hidayat, W. (2017). Ekstraksi Kata Dasar Secara Berjenjang ( Incremental Stemming ) Berbasis Aturan Morfologi untuk Teks Berbahasa Indonesia, 9(2), 166–171.
- Karim, M. A., Kaykobad, M., & Murshed, M. (2013). *Technical Challenges and Design Issues in Bangla Language Processing*. USA: Information Science Reference (an imprint of IGI Global).
- Kemendikbud, T. P. (2015). *Rencana Strategis Badan Pengembangan dan Pembinaan Bahasa 2015-2019*. Badan Pengembangan dan Pembinaan Bahasa Kemendikbud.
- Kumar, E. (2011). *Natural Language Processing*. New Delhi: I.K. International Publishing House Pvt. Ltd.
- Levitin, A. (2003). *Introduction to The Design and Analysis of Algorithms*. Boston: Addison Wesley.
- Maulidi, R. (2016). Stemmer untuk Bahasa Madura dengan Modifikasi Metode Enhanced Confix Stripping Stemmer. *ResearchGate*, (December), 12–15.
- Munir, R. (2011). *Algoritma & Pemrograman Dalam Bahasa Pascal dan C (Edisi Revisi)*. Bandung: Informatika Bandung.
- Nopiyanti, D., & Sekarwati, K. A. (2014). Aplikasi Pencarian Kata Dasar Dokumen Berbahasa Indonesia dengan Metode Stemming Porter Menggunakan PHP & MYSQL, 8(Kommit), 215–222.
- Paice, C. D. (1996). Method for eveluation of stemming algorithms based on error counting. *Journal of the American Society for Information Science*, 47(8).
- Schofield, A., & Mimno, D. (2016). Comparing Apples to Apple : The Effects of Stemmers on Topic Models. *Association for Computational Linguistics*, 4, 287–300.
- Setiawan, R., Kurniawan, A., Budiharto, W., Kartowisastro, I. H., & Prabowo, H. (2016). Flexible Affix Classification for Stemming Indonesian Language. *IEEE*.
- Sharma, D. (2012). Stemming Algorithms : A Comparative Study and their Analysis, 4(3), 7–12.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- Soeherman, B., & Pinontoan, M. (2008). *Designing Information System*. Jakarta: PT. Elex Media Komputindo.
- Sulhan, M., & Kurniawan, R. (2014). METODE STEMMING SEBAGAI PREPROCESSING PADA FILTER KATA PORNO MELALUI ASPEK PENDIDIKAN, 52–60.
- Sutojo, T., Mulyanto, E., & Suhartono, V. (2011). *Kecerdasan Buatan*. Yogyakarta: Penerbit ANDI.
- Wedhawati, Nurlina, W. E. S., Setiyanto, E., & Sukesti, R. (2006). *Tata Bahasa Jawa Mutakhir*. Yogyakarta: Penerbit Kanisius(Anggota IKAPI).
- Wijaya, T. (2005). *Juaro*. Lampung: Rosa.
- Winarti, T., Kerami, D., Etp, L., & Sudiro, S. A. (2017). Improving Stemming Algorithm Using Morphological Rules. *International Journal on Advanced Science Engineering Information Technology*, 7(5), 1758–1764. <https://doi.org/10.18517/ijaseit.7.5.1705>
- Yasiroh, N. (2013). *Proses Morfologis Bahasa Melayu Palembang*. Universitas Negeri Yogyakarta.
- Zaenal, A., & Junaiyah, H. (2007). *Morfologi bentuk, makna dan fungsi*. Jakarta: PT. Grasindo.
- Zulfikar, A. F. (2017). Pengembangan Algoritma Stemming Bahasa Indonesia dengan Pendekatan Dictionary Base Stemming untuk Menentukan Kata. *Jurnal Informatika*, 2.



## LAMPIRAN A

### DATA KATA UJI DAN HASIL STEMMING

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
1	aba	aba	aba	Benar	Ya	Ada
2	abab	abab	abab	Benar	Ya	Ada
3	ababmu	abab	abab	Benar	Ya	Ada
4	abanyo	aba	aba	Benar	Ya	Ada
5	ada	ada	ada	Benar	Ya	Ada
6	adik	adik	adik	Benar	Ya	Ada
7	adiknyo	adik	adik	Benar	Ya	Ada
8	adil	adil	adil	Benar	Ya	Ada
9	ado	ado	ado	Benar	Ya	Ada
10	aduh	aduh	aduh	Benar	Ya	Ada
11	akhirnyo	akhir	akhir	Benar	Ya	Ada
12	ajaibnyo	ajaib	ajaib	Benar	Ya	Ada
13	akan	akan	akan	Benar	Ya	Ada
14	aku	aku	aku	Benar	Ya	Ada
15	akui	aku	aku	Benar	Ya	Ada
16	amaknyo	amak	amak	Benar	Ya	Ada
17	ampun	ampun	ampun	Benar	Ya	Ada
18	anak	anak	anak	Benar	Ya	Ada
19	anakku	anak	anak	Benar	Ya	Ada
20	anaknyo	anak	anak	Benar	Ya	Ada
21	angin	angin	angin	Benar	Ya	Ada
22	angkatnyo	angkat	angkat	Benar	Ya	Ada
23	anterke	anter	anter	Benar	Ya	Ada
24	api	api	api	Benar	Ya	Ada
25	apo	apo	apo	Benar	Ya	Ada
26	aponyo	apo	apo	Benar	Ya	Ada
27	ari	ari	ari	Benar	Ya	Ada
28	asal	asal	asal	Benar	Ya	Ada
29	ati	ati	ati	Benar	Ya	Ada
30	atinyo	ati	ati	Benar	Ya	Ada
31	awak	awak	awak	Benar	Ya	Ada
32	ayuknyo	ayuk	ayuk	Benar	Ya	Ada
33	badan	badan	badan	Benar	Ya	Ada
34	bae	bae	bae	Benar	Ya	Ada
35	baek	baek	baek	Benar	Ya	Ada
36	bahayo	bahayo	bahayo	Benar	Ya	Ada
37	baikla	baik	baik	Benar	Ya	Ada
38	baju	baju	baju	Benar	Ya	Ada

#### Hak Cipta Dilindungi Undang-Undang

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
- Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
39	bajunyo	baju	baju	Benar	Ya	Ada
40	balek	balek	balek	Benar	Ya	Ada
41	balik	balik	balik	Benar	Ya	Ada
42	balikla	balik	balik	Benar	Ya	Ada
43	bandingnyo	banding	banding	Benar	Ya	Ada
44	bandit	bandit	bandit	Benar	Ya	Ada
45	bantahan	bantah	bantah	Benar	Ya	Ada
46	banyak	banyak	banyak	Benar	Ya	Ada
47	banyu	banyu	banyu	Benar	Ya	Ada
48	bapaknyo	bapak	bapak	Benar	Ya	Ada
49	barang	barang	barang	Benar	Ya	Ada
50	barengan	bareng	bareng	Benar	Ya	Ada
51	baru	baru	baru	Benar	Ya	Ada
52	basa	basa	basa	Benar	Ya	Ada
53	batang	batang	batang	Benar	Ya	Ada
54	batu	batu	batu	Benar	Ya	Ada
55	bawa	bawa	bawa	Benar	Ya	Ada
56	bawaannyo	bawa	bawa	Benar	Ya	Ada
57	bawang	bawang	bawang	Benar	Ya	Ada
58	bayar	bayar	bayar	Benar	Ya	Ada
59	bebua	bua	bua	Benar	Ya	Ada
60	bedayo	dayo	dayo	Benar	Ya	Ada
61	bedoa	doa	doa	Benar	Ya	Ada
62	begawe	gawe	gawe	Benar	Ya	Ada
63	begawi	gawi	gawi	Benar	Ya	Ada
64	begoco	goco	goco	Benar	Ya	Ada
65	bejajo	bejajo	bejajo	Benar	Ya	Ada
66	bejalan	bejalan	bejalan	Benar	Ya	Ada
67	bekacolah	kaco	kaco	Benar	Ya	Ada
68	bela	bela	bela	Benar	Ya	Ada
69	belalang	belalang	belalang	Benar	Ya	Ada
70	beli	beli	beli	Benar	Ya	Ada
71	belum	belum	belum	Benar	Ya	Ada
72	bemalu	malu	malu	Benar	Ya	Ada
73	benar	benar	benar	Benar	Ya	Ada
74	benci	benci	benci	Benar	Ya	Ada
75	bener	bener	bener	Benar	Ya	Ada
76	bepesen	pesen	pesen	Benar	Ya	Ada
77	bepisa	pisa	pisa	Benar	Ya	Ada
78	beradik	adik	adik	Benar	Ya	Ada
79	beranak	beranak	beranak	Benar	Ya	Ada
80	beranggepan	anggep	anggep	Benar	Ya	Ada
81	berangkat	berangkat	berangkat	Benar	Ya	Ada

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
82	berapo	berapo	berapo	Benar	Ya	Ada
83	beras	beras	beras	Benar	Ya	Ada
84	berat	berat	berat	Benar	Ya	Ada
85	beratap	atap	atap	Benar	Ya	Ada
86	bercerai	cerai	cerai	Benar	Ya	Ada
87	beringkesla	ringkes	ringkes	Benar	Ya	Ada
88	bertahan	tahan	tahan	Benar	Ya	Ada
89	bertangkup	tangkup	tangkup	Benar	Ya	Ada
90	beruju	beruju	beruju	Benar	Ya	Ada
91	berupo	rupo	rupo	Benar	Ya	Ada
92	besak	besak	besak	Benar	Ya	Ada
93	besaknyo	besak	besak	Benar	Ya	Ada
94	besok	besok	besok	Benar	Ya	Ada
95	besuk	besuk	besuk	Benar	Ya	Ada
96	betamba	tamba	tamba	Benar	Ya	Ada
97	betangkup	tangkup	tangkup	Benar	Ya	Ada
98	betanyo	tanyo	tanyo	Benar	Ya	Ada
99	betemu	temu	temu	Benar	Ya	Ada
100	betino	betino	betino	Benar	Ya	Ada
101	biar	biar	biar	Benar	Ya	Ada
102	biarlah	biar	biar	Benar	Ya	Ada
103	biasonyo	biaso	biaso	Benar	Ya	Ada
104	bibir	bibir	bibir	Benar	Ya	Ada
105	bicek	bicek	bicek	Benar	Ya	Ada
106	bicik	bicik	bicik	Benar	Ya	Ada
107	bila	bila	bila	Benar	Ya	Ada
108	bilang	bilang	bilang	Benar	Ya	Ada
109	bini	bini	bini	Benar	Ya	Ada
110	<b>bininyo</b>	<b>ini</b>	<b>bini</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
111	biso	biso	biso	Benar	Ya	Ada
112	bole	bole	bole	Benar	Ya	Ada
113	bolela	ole	ole	Benar	Ya	Ada
114	buanyo	bua	bua	Benar	Ya	Ada
115	buat	buat	buat	Benar	Ya	Ada
116	budak	budak	budak	Benar	Ya	Ada
117	budaknyo	budak	budak	Benar	Ya	Ada
118	buka	buka	buka	Benar	Ya	Ada
119	bukan	bukan	bukan	Benar	Ya	Ada
120	buku	buku	buku	Benar	Ya	Ada
121	bulan	bulan	bulan	Benar	Ya	Ada
122	buruk	buruk	buruk	Benar	Ya	Ada
123	buyan	buyan	buyan	Benar	Ya	Ada
124	cak	cak	cak	Benar	Ya	Ada



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
125	calaknyo	calak	calak	Benar	Ya	Ada
126	cari	cari	cari	Benar	Ya	Ada
127	carike	cari	cari	Benar	Ya	Ada
128	caro	caro	caro	Benar	Ya	Ada
129	celanamu	celana	celana	Benar	Ya	Ada
130	cemburuku	cemburu	cemburu	Benar	Ya	Ada
131	cepat	cepat	cepat	Benar	Ya	Ada
132	cerai	cerai	cerai	Benar	Ya	Ada
133	cerito	cerito	cerito	Benar	Ya	Ada
134	ceritoke	cerito	cerito	Benar	Ya	Ada
135	ceritonyo	cerito	cerito	Benar	Ya	Ada
136	cindo	cindo	cindo	Benar	Ya	Ada
137	coba	coba	coba	Benar	Ya	Ada
138	cobo	cobo	cobo	Benar	Ya	Ada
139	cocok	cocok	cocok	Benar	Ya	Ada
140	comot	comot	comot	Benar	Ya	Ada
141	cubo	cubo	cubo	Benar	Ya	Ada
142	cugak	cugak	cugak	Benar	Ya	Ada
143	cumen	cumen	cumen	Benar	Ya	Ada
144	dagang	dagang	dagang	Benar	Ya	Ada
145	dagangan	dagang	dagang	Benar	Ya	Ada
146	dagangannyo	dagang	dagang	Benar	Ya	Ada
147	dak	dak	dak	Benar	Ya	Ada
148	dakke	dak	dak	Benar	Ya	Ada
149	dan	dan	dan	Benar	Ya	Ada
150	dapat	dapat	dapat	Benar	Ya	Ada
151	dapet	dapet	dapet	Benar	Ya	Ada
152	dapetke	dapet	dapet	Benar	Ya	Ada
153	darat	darat	darat	Benar	Ya	Ada
154	dari	dari	dari	Benar	Ya	Ada
155	darimano	darimano	darimano	Benar	Ya	Ada
156	daripado	daripado	daripado	Benar	Ya	Ada
157	datang	datang	datang	Benar	Ya	Ada
158	datangla	datang	datang	Benar	Ya	Ada
159	dengan	dengan	dengan	Benar	Ya	Ada
160	dengen	dengen	dengen	Benar	Ya	Ada
161	denget	denget	denget	Benar	Ya	Ada
162	dewek	dewek	dewek	Benar	Ya	Ada
163	diampirkenyola	ampir	ampir	Benar	Ya	Ada
164	dianggepnyo	anggep	anggep	Benar	Ya	Ada
165	dibayarnya	bayar	bayar	Benar	Ya	Ada
166	dibelinyo	beli	beli	Benar	Ya	Ada
167	dibesakke	besak	besak	Benar	Ya	Ada

# Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
168	dicandaknyo	candak	candak	Benar	Ya	Ada
169	diceritokenyola	cerito	cerito	Benar	Ya	Ada
170	didiami	diem	diem	Benar	Ya	Ada
171	diembeknyola	embek	embek	Benar	Ya	Ada
172	digaweke	gawe	gawe	Benar	Ya	Ada
173	digendong	gendong	gendong	Benar	Ya	Ada
174	digusur	gusur	gusur	Benar	Ya	Ada
175	diinjak	injak	injak	Benar	Ya	Ada
176	dijadikenyola	jadi	jadi	Benar	Ya	Ada
177	dijalaninyola	jalan	jalan	Benar	Ya	Ada
178	dijanji	janji	janji	Benar	Ya	Ada
179	dijingok	jingok	jingok	Benar	Ya	Ada
180	dijingoknyola	jingok	jingok	Benar	Ya	Ada
181	dijualnyola	jual	jual	Benar	Ya	Ada
182	dikatoke	kato	kato	Benar	Ya	Ada
183	dikayukenyola	kayu	kayu	Benar	Ya	Ada
184	dikejar	kejar	kejar	Benar	Ya	Ada
185	dikongkon	kongkon	kongkon	Benar	Ya	Ada
186	<b>dikongkonyola</b>	<b>dikongkonyola</b>	<b>kongkon</b>	<b>Salah</b>	<b>Tidak</b>	<b>Tidak</b>
187	dilamar	lamar	lamar	Benar	Ya	Ada
188	dilamarnya	lamar	lamar	Benar	Ya	Ada
189	dilepaskenyola	lepas	lepas	Benar	Ya	Ada
190	<b>dimakan</b>	<b>mak</b>	<b>makan</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
191	<b>dimakannyola</b>	<b>mak</b>	<b>makan</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
192	dinamoi	namo	namo	Benar	Ya	Ada
193	dio	dio	dio	Benar	Ya	Ada
194	diola	ola	ola	Benar	Ya	Ada
195	diomongkenyola	omong	omong	Benar	Ya	Ada
196	dipecayo	pecayo	pecayo	Benar	Ya	Ada
197	dipeteknyola	petek	petek	Benar	Ya	Ada
198	dipikirke	pikir	pikir	Benar	Ya	Ada
199	dirinyola	diri	diri	Benar	Ya	Ada
200	disambersnyola	samber	samber	Benar	Ya	Ada
201	disano	sano	sano	Benar	Ya	Ada
202	disebut	sebut	sebut	Benar	Ya	Ada
203	disertoi	serto	serto	Benar	Ya	Ada
204	disimbati	simbat	simbat	Benar	Ya	Ada
205	ditakuti	takut	takut	Benar	Ya	Ada
206	ditanemnyola	tanem	tanem	Benar	Ya	Ada
207	ditangkep	tangkep	tangkep	Benar	Ya	Ada
208	ditangkepnyola	tangkep	tangkep	Benar	Ya	Ada
209	ditembak	tembak	tembak	Benar	Ya	Ada
210	diterimonyola	terimo	terimo	Benar	Ya	Ada

[illegible]

- | No  | Kata Uji         | Stemming Algoritma | Stemming Validator | Hasil Sistem | Sesuai Harapan | Kamus |
|-----|------------------|--------------------|--------------------|--------------|----------------|-------|
| 211 | diteruskenyola   | terus              | terus              | Benar        | Ya             | Ada   |
| 212 | ditindih         | tindih             | tindih             | Benar        | Ya             | Ada   |
| 213 | ditinggalke      | tinggal            | tinggal            | Benar        | Ya             | Ada   |
| 214 | ditinggalkenyo   | tinggal            | tinggal            | Benar        | Ya             | Ada   |
| 215 | ditinggalkenyola | tinggal            | tinggal            | Benar        | Ya             | Ada   |
| 216 | diturunke        | turun              | turun              | Benar        | Ya             | Ada   |
| 217 | diurus           | urus               | urus               | Benar        | Ya             | Ada   |
| 218 | dusirnyo         | usir               | usir               | Benar        | Ya             | Ada   |
| 219 | duit             | duit               | duit               | Benar        | Ya             | Ada   |
| 220 | duitnyo          | duit               | duit               | Benar        | Ya             | Ada   |
| 221 | dulu             | dulu               | dulu               | Benar        | Ya             | Ada   |
| 222 | dulur            | dulur              | dulur              | Benar        | Ya             | Ada   |
| 223 | duo              | duo                | duo                | Benar        | Ya             | Ada   |
| 224 | dusonyo          | duso               | duso               | Benar        | Ya             | Ada   |
| 225 | embekla          | embek              | embek              | Benar        | Ya             | Ada   |
| 226 | gadisnyo         | gadis              | gadis              | Benar        | Ya             | Ada   |
| 227 | gaib             | gaib               | gaib               | Benar        | Ya             | Ada   |
| 228 | gaji             | gaji               | gaji               | Benar        | Ya             | Ada   |
| 229 | gajiku           | gaji               | gaji               | Benar        | Ya             | Ada   |
| 230 | galak            | galak              | galak              | Benar        | Ya             | Ada   |
| 231 | galo             | galo               | galo               | Benar        | Ya             | Ada   |
| 232 | galonyo          | galonyo            | galonyo            | Benar        | Ya             | Ada   |
| 233 | gampang          | gampang            | gampang            | Benar        | Ya             | Ada   |
| 234 | ganti            | ganti              | ganti              | Benar        | Ya             | Ada   |
| 235 | gaweke           | gawe               | gawe               | Benar        | Ya             | Ada   |
| 236 | gawinyo          | gawi               | gawi               | Benar        | Ya             | Ada   |
| 237 | gelap            | gelap              | gelap              | Benar        | Ya             | Ada   |
| 238 | gembironyo       | gembiro            | gembiro            | Benar        | Ya             | Ada   |
| 239 | giliran          | gilir              | gilir              | Benar        | Ya             | Ada   |
| 240 | gilo             | gilo               | gilo               | Benar        | Ya             | Ada   |
| 241 | guno             | guno               | guno               | Benar        | Ya             | Ada   |
| 242 | guru             | guru               | guru               | Benar        | Ya             | Ada   |
| 243 | habis            | habis              | habis              | Benar        | Ya             | Ada   |
| 244 | hadam            | hadam              | hadam              | Benar        | Ya             | Ada   |
| 245 | hanya            | hanya              | hanya              | Benar        | Ya             | Ada   |
| 246 | harus            | harus              | harus              | Benar        | Ya             | Ada   |
| 247 | hebat            | hebat              | hebat              | Benar        | Ya             | Ada   |
| 248 | ibo              | ibo                | ibo                | Benar        | Ya             | Ada   |
| 249 | idak             | idak               | idak               | Benar        | Ya             | Ada   |
| 250 | idup             | idup               | idup               | Benar        | Ya             | Ada   |
| 251 | idupla           | idup               | idup               | Benar        | Ya             | Ada   |
| 252 | ikhlas           | ikhlas             | ikhlas             | Benar        | Ya             | Ada   |
| 253 | ingeti           | inget              | inget              | Benar        | Ya             | Ada   |



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
254	ini	ini	ini	Benar	Ya	Ada
255	intinyo	inti	inti	Benar	Ya	Ada
256	isap	isap	isap	Benar	Ya	Ada
257	itu	itu	itu	Benar	Ya	Ada
258	itula	itu	itu	Benar	Ya	Ada
259	itulah	itu	itu	Benar	Ya	Ada
260	jadi	jadi	jadi	Benar	Ya	Ada
261	jadilah	jadi	jadi	Benar	Ya	Ada
262	jago	jago	jago	Benar	Ya	Ada
263	jagola	jago	jago	Benar	Ya	Ada
264	jahat	jahat	jahat	Benar	Ya	Ada
265	jalan	jalan	jalan	Benar	Ya	Ada
266	jalur	jalur	jalur	Benar	Ya	Ada
267	jam	jam	jam	Benar	Ya	Ada
268	jaman	jaman	jaman	Benar	Ya	Ada
269	jamin	jamin	jamin	Benar	Ya	Ada
270	jando	jando	jando	Benar	Ya	Ada
271	jangan	jangan	jangan	Benar	Ya	Ada
272	janganla	jangan	jangan	Benar	Ya	Ada
273	janganlah	jangan	jangan	Benar	Ya	Ada
274	jao	jao	jao	Benar	Ya	Ada
275	jeragan	jeragan	jeragan	Benar	Ya	Ada
276	jeragannyo	jeragan	jeragan	Benar	Ya	Ada
277	jeru	jeru	jeru	Benar	Ya	Ada
278	jingok	jingok	jingok	Benar	Ya	Ada
279	jingoknyo	jingok	jingok	Benar	Ya	Ada
280	jualan	jual	jual	Benar	Ya	Ada
281	juaro	juaro	juaro	Benar	Ya	Ada
282	jugo	jugo	jugo	Benar	Ya	Ada
283	kabar	kabar	kabar	Benar	Ya	Ada
284	kagek	kagek	kagek	Benar	Ya	Ada
285	kalian	kalian	kalian	Benar	Ya	Ada
286	kalu	kalu	kalu	Benar	Ya	Ada
287	kamek	kamek	kamek	Benar	Ya	Ada
288	kami	kami	kami	Benar	Ya	Ada
289	kampung	kampung	kampung	Benar	Ya	Ada
290	kantor	kantor	kantor	Benar	Ya	Ada
291	kanyut	kanyut	kanyut	Benar	Ya	Ada
292	kanyutnyo	kanyut	kanyut	Benar	Ya	Ada
293	kapak	kapak	kapak	Benar	Ya	Ada
294	kapal	kapal	kapal	Benar	Ya	Ada
295	kapalnyo	kapal	kapal	Benar	Ya	Ada
296	kartunyo	kartu	kartu	Benar	Ya	Ada



# Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
297	kasi	kasi	kasi	Benar	Ya	Ada
298	kasinyo	kasi	kasi	Benar	Ya	Ada
299	katik	katik	katik	Benar	Ya	Ada
300	katokela	kato	kato	Benar	Ya	Ada
301	katonyo	kato	kato	Benar	Ya	Ada
302	kau	kau	kau	Benar	Ya	Ada
303	kawan	kawan	kawan	Benar	Ya	Ada
304	kawannyo	kawan	kawan	Benar	Ya	Ada
305	kawin	kawin	kawin	Benar	Ya	Ada
306	kayo	kayo	kayo	Benar	Ya	Ada
307	kayu	kayu	kayu	Benar	Ya	Ada
308	keadaannya	ada	ada	Benar	Ya	Ada
309	kebenaran	benar	benar	Benar	Ya	Ada
310	kecik	kecik	kecik	Benar	Ya	Ada
311	kedengen	dengen	dengen	Benar	Ya	Ada
312	keidupan	idup	idup	Benar	Ya	Ada
313	kejujuran	jujur	jujur	Benar	Ya	Ada
314	kekuasaan	kuasa	kuasa	Benar	Ya	Ada
315	kelakar	kelakar	kelakar	Benar	Ya	Ada
316	kelaso	kelaso	kelaso	Benar	Ya	Ada
317	keliwat	keliwat	keliwat	Benar	Ya	Ada
318	keluar	luar	luar	Benar	Ya	Ada
319	keluargo	keluargo	keluargo	Benar	Ya	Ada
320	kemano	kemano	mano	Benar	Tidak	Ada
321	kemanoke	mano	mano	Benar	Ya	Ada
322	kemudi	kemudi	kemudi	Benar	Ya	Ada
323	kemudinyo	kemudi	kemudi	Benar	Ya	Ada
324	keno	keno	keno	Benar	Ya	Ada
325	kepada	pado	pado	Benar	Ya	Ada
326	kering	kering	kering	Benar	Ya	Ada
327	kerno	kerno	kerno	Benar	Ya	Ada
328	kerumanyo	ruma	ruma	Benar	Ya	Ada
329	kesel	kesel	kesel	Benar	Ya	Ada
330	kesiangan	siang	siang	Benar	Ya	Ada
331	kesukoan	suko	suko	Benar	Ya	Ada
332	kesulitan	sulit	sulit	Benar	Ya	Ada
333	kesultanan	sultan	sultan	Benar	Ya	Ada
334	ketemu	temu	temu	Benar	Ya	Ada
335	khawatir	khawatir	khawatir	Benar	Ya	Ada
336	kito	kito	kito	Benar	Ya	Ada
337	kodak	kodak	kodak	Benar	Ya	Ada
338	kuaduke	adu	adu	Benar	Ya	Ada
339	kuatir	kuatir	kuatir	Benar	Ya	Ada

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
340	kujadike	jadi	jadi	Benar	Ya	Ada
341	kujamin	jamin	jamin	Benar	Ya	Ada
342	kulupoke	lupo	lupo	Benar	Ya	Ada
343	kupindahke	pindah	pindah	Benar	Ya	Ada
344	kuping	kuping	kuping	Benar	Ya	Ada
345	lading	lading	lading	Benar	Ya	Ada
346	ladingnyo	lading	lading	Benar	Ya	Ada
347	lagi	lagi	lagi	Benar	Ya	Ada
348	lahir	lahir	lahir	Benar	Ya	Ada
349	lain	lain	lain	Benar	Ya	Ada
350	lainnyo	lain	lain	Benar	Ya	Ada
351	lamo	lamo	lamo	Benar	Ya	Ada
352	lanang	lanang	lanang	Benar	Ya	Ada
353	langgar	langgar	langgar	Benar	Ya	Ada
354	langsung	langsung	langsung	Benar	Ya	Ada
355	lapar	lapar	lapar	Benar	Ya	Ada
356	lawan	lawan	lawan	Benar	Ya	Ada
357	lebi	lebi	lebi	Benar	Ya	Ada
358	lebiannyo	lebi	lebi	Benar	Ya	Ada
359	leket	leket	leket	Benar	Ya	Ada
360	lemak	lemak	lemak	Benar	Ya	Ada
361	lemaknyo	lemak	lemak	Benar	Ya	Ada
362	maaf	maaf	maaf	Benar	Ya	Ada
363	maen	maen	maen	Benar	Ya	Ada
364	mahal	mahal	mahal	Benar	Ya	Ada
365	mahasiswa	mahasiswa	mahasiswa	Benar	Ya	Ada
366	main	main	main	Benar	Ya	Ada
367	mainan	main	main	Benar	Ya	Ada
368	maju	maju	maju	Benar	Ya	Ada
369	mak	mak	mak	Benar	Ya	Ada
370	makan	makan	makan	Benar	Ya	Ada
371	makanan	akan	akan	Benar	Ya	Ada
372	makanlah	akan	akan	Benar	Ya	Ada
373	make	pake	pake	Benar	Ya	Ada
374	maki	maki	maki	Benar	Ya	Ada
375	makin	makin	makin	Benar	Ya	Ada
376	maknyo	bak	bak	Benar	Ya	Ada
377	makso	makso	makso	Benar	Ya	Ada
378	maksudnyo	maksud	maksud	Benar	Ya	Ada
379	malam	malam	malam	Benar	Ya	Ada
380	malu	malu	malu	Benar	Ya	Ada
381	maluke	malu	malu	Benar	Ya	Ada
382	mamang	mamang	mamang	Benar	Ya	Ada



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
383	manis	manis	manis	Benar	Ya	Ada
384	mano	mano	mano	Benar	Ya	Ada
385	manusio	manusio	manusio	Benar	Ya	Ada
386	maranyo	mara	mara	Benar	Ya	Ada
387	masak	masak	masak	Benar	Ya	Ada
388	masi	masi	masi	Benar	Ya	Ada
389	masih	masih	masih	Benar	Ya	Ada
390	masok	masok	masok	Benar	Ya	Ada
391	masokla	masok	masok	Benar	Ya	Ada
392	mata	pata	pata	Benar	Ya	Ada
393	matek	matek	matek	Benar	Ya	Ada
394	mati	mati	mati	Benar	Ya	Ada
395	mato	mato	mato	Benar	Ya	Ada
396	mecak	pecak	pecak	Benar	Ya	Ada
397	mela	bela	bela	Benar	Ya	Ada
398	melarat	melarat	melarat	Benar	Ya	Ada
399	melaratnyo	larat	larat	Benar	Ya	Ada
400	meleng	meleng	meleng	Benar	Ya	Ada
401	memaluke	malu	malu	Benar	Ya	Ada
402	memang	memang	memang	Benar	Ya	Ada
403	menang	menang	menang	Benar	Ya	Ada
404	mendep	mendep	mendep	Benar	Ya	Ada
405	menerimo	terimo	terimo	Benar	Ya	Ada
406	menerimonyo	terimo	terimo	Benar	Ya	Ada
407	mengajaknyo	ajak	ajak	Benar	Ya	Ada
408	mengamuk	amuk	amuk	Benar	Ya	Ada
409	merajok	merajok	merajok	Benar	Ya	Ada
410	metu	metu	metu	Benar	Ya	Ada
411	minta	minta	minta	Benar	Ya	Ada
412	minyak	minyak	minyak	Benar	Ya	Ada
413	miskin	miskin	miskin	Benar	Ya	Ada
414	modal	modal	modal	Benar	Ya	Ada
415	mudo	mudo	mudo	Benar	Ya	Ada
416	mulut	mulut	mulut	Benar	Ya	Ada
417	mungkin	mungkin	mungkin	Benar	Ya	Ada
418	mutung	mutung	mutung	Benar	Ya	Ada
419	mutuskenyo	putus	putus	Benar	Ya	Ada
420	naik	naik	naik	Benar	Ya	Ada
421	nak	nak	nak	Benar	Ya	Ada
422	namoku	namo	namo	Benar	Ya	Ada
423	nangkep	tangkep	tangkep	Benar	Ya	Ada
424	nantang	tantang	tantang	Benar	Ya	Ada
425	napas	napas	napas	Benar	Ya	Ada

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
426	negeri	negeri	negeri	Benar	Ya	Ada
427	nemui	temu	temu	Benar	Ya	Ada
428	nenger	denger	denger	Benar	Ya	Ada
429	neruske	erus	erus	Benar	Ya	Ada
430	ngajak	ajak	ajak	Benar	Ya	Ada
431	ngakuinyo	aku	aku	Benar	Ya	Ada
432	ngamuk	amuk	amuk	Benar	Ya	Ada
433	ngapak	kapak	kapak	Benar	Ya	Ada
434	ngapo	ngapo	ngapo	Benar	Ya	Ada
435	ngarepke	arep	arep	Benar	Ya	Ada
436	ngatoi	kato	kato	Benar	Ya	Ada
437	ngatokenyo	kato	kato	Benar	Ya	Ada
438	ngawak	ngawak	ngawak	Benar	Ya	Ada
439	ngejer	jejer	jejer	Benar	Ya	Ada
440	ngelamar	lamar	lamar	Benar	Ya	Ada
441	ngendong	gendong	gendong	Benar	Ya	Ada
442	ngengjuk	ngengjuk	ngengjuk	Benar	Ya	Ada
443	ngenoi	keno	keno	Benar	Ya	Ada
444	ngingok	jingok	jingok	Benar	Ya	Ada
445	ngomong	ngomong	ngomong	Benar	Ya	Ada
446	<b>ngopo</b>	<b>gopo</b>	<b>opo</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
447	ngrewangi	rewang	rewang	Benar	Ya	Ada
448	ngucapke	ucap	ucap	Benar	Ya	Ada
449	ngumpuli	kumpul	kumpul	Benar	Ya	Ada
450	ngumpulke	kumpul	kumpul	Benar	Ya	Ada
451	ngurangi	kurang	kurang	Benar	Ya	Ada
452	nian	nian	nian	Benar	Ya	Ada
453	nidukke	tiduk	tiduk	Benar	Ya	Ada
454	niemi	diem	diem	Benar	Ya	Ada
455	nuju	tuju	tuju	Benar	Ya	Ada
456	nulung	tulung	tulung	Benar	Ya	Ada
457	nunggu	tunggu	tunggu	Benar	Ya	Ada
458	nyago	jago	jago	Benar	Ya	Ada
459	nyagoi	jago	jago	Benar	Ya	Ada
460	nyandak	candak	candak	Benar	Ya	Ada
461	nyari	cari	cari	Benar	Ya	Ada
462	nyatoke	kato	kato	Benar	Ya	Ada
463	nyawo	nyawo	nyawo	Benar	Ya	Ada
464	nyicipnyo	cicip	cicip	Benar	Ya	Ada
465	nyingok	nyingok	nyingok	Benar	Ya	Ada
466	nyucukke	cucuk	cucuk	Benar	Ya	Ada
467	nyusake	susa	susa	Benar	Ya	Ada
468	ole	ole	ole	Benar	Ya	Ada

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
469	omongan	omong	omong	Benar	Ya	Ada
470	pacak	pacak	pacak	Benar	Ya	Ada
471	pado	pado	pado	Benar	Ya	Ada
472	paginyo	pagi	pagi	Benar	Ya	Ada
473	pakela	pake	pake	Benar	Ya	Ada
474	panas	panas	panas	Benar	Ya	Ada
475	panik	panik	panik	Benar	Ya	Ada
476	parak	parak	parak	Benar	Ya	Ada
477	pasti	pasti	pasti	Benar	Ya	Ada
478	pata	pata	pata	Benar	Ya	Ada
479	payola	payo	payo	Benar	Ya	Ada
480	peca	peca	peca	Benar	Ya	Ada
481	pecak	pecak	pecak	Benar	Ya	Ada
482	<b>pecayola</b>	<b>cayo</b>	<b>pecayo</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
483	pegi	pegi	pegi	Benar	Ya	Ada
484	pegila	pegi	pegi	Benar	Ya	Ada
485	pekikke	pekik	pekik	Benar	Ya	Ada
486	pelembang	pelembang	pelembang	Benar	Ya	Ada
487	pencariannyo	cari	cari	Benar	Ya	Ada
488	pengabisan	abis	abis	Benar	Ya	Ada
489	pengadilan	adil	adil	Benar	Ya	Ada
490	<b>pengayunyo</b>	<b>ayu</b>	<b>kayu</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
491	penjagonyo	jago	jago	Benar	Ya	Ada
492	penjaro	penjaro	penjaro	Benar	Ya	Ada
493	pentung	pentung	pentung	Benar	Ya	Ada
494	penyarian	cari	cari	Benar	Ya	Ada
495	peranti	peranti	peranti	Benar	Ya	Ada
496	perasaan	perasaan	perasaan	Benar	Ya	Ada
497	perau	perau	perau	Benar	Ya	Ada
498	perhatian	perhatian	perhatian	Benar	Ya	Ada
499	petir	petir	petir	Benar	Ya	Ada
500	pikirnyo	pikir	pikir	Benar	Ya	Ada
501	pilu	pilu	pilu	Benar	Ya	Ada
502	pinggir	pinggir	pinggir	Benar	Ya	Ada
503	pinggiran	pinggir	pinggir	Benar	Ya	Ada
504	pokoknyo	pokok	pokok	Benar	Ya	Ada
505	polisi	polisi	polisi	Benar	Ya	Ada
506	pondok	pondok	pondok	Benar	Ya	Ada
507	pondoknyo	pondok	pondok	Benar	Ya	Ada
508	puas	puas	puas	Benar	Ya	Ada
509	pucuk	pucuk	pucuk	Benar	Ya	Ada
510	pulo	pulo	pulo	Benar	Ya	Ada
511	punyo	punyo	punyo	Benar	Ya	Ada



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
512	putihnyo	putih	putih	Benar	Ya	Ada
513	ragu	ragu	ragu	Benar	Ya	Ada
514	rai	rai	rai	Benar	Ya	Ada
515	rainyo	rai	rai	Benar	Ya	Ada
516	rambut	rambut	rambut	Benar	Ya	Ada
517	rambutnyo	rambut	rambut	Benar	Ya	Ada
518	rami	rami	rami	Benar	Ya	Ada
519	rampok	rampok	rampok	Benar	Ya	Ada
520	rasannyo	rasan	rasan	Benar	Ya	Ada
521	raso	raso	raso	Benar	Ya	Ada
522	rasoke	raso	raso	Benar	Ya	Ada
523	rasonyo	raso	raso	Benar	Ya	Ada
524	regonyo	rego	rego	Benar	Ya	Ada
525	ribut	ribut	ribut	Benar	Ya	Ada
526	ruma	ruma	ruma	Benar	Ya	Ada
527	rumanyo	ruma	ruma	Benar	Ya	Ada
528	ruponyo	rupo	rupo	Benar	Ya	Ada
529	ruso	ruso	ruso	Benar	Ya	Ada
530	sakit	sakit	sakit	Benar	Ya	Ada
531	sakitnyo	sakit	sakit	Benar	Ya	Ada
532	sambil	sambil	sambil	Benar	Ya	Ada
533	samo	samo	samo	Benar	Ya	Ada
534	sampai	sampai	sampai	Benar	Ya	Ada
535	sampela	sampe	sampe	Benar	Ya	Ada
536	sampi	sampi	sampi	Benar	Ya	Ada
537	sampila	sampi	sampi	Benar	Ya	Ada
538	sanak	sanak	sanak	Benar	Ya	Ada
539	sano	sano	sano	Benar	Ya	Ada
540	sapi	sapi	sapi	Benar	Ya	Ada
541	sari	ari	ari	Benar	Ya	Ada
542	saro	saro	saro	Benar	Ya	Ada
543	satu	satu	satu	Benar	Ya	Ada
544	sayang	sayang	sayang	Benar	Ya	Ada
545	sayangnyo	sayang	sayang	Benar	Ya	Ada
546	sebasingan	basing	basing	Benar	Ya	Ada
547	<b>sebelum</b>	<b>sebelum</b>	<b>belum</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
548	sebelumnyo	belum	belum	Benar	Ya	Ada
549	sebenarnya	benar	benar	Benar	Ya	Ada
550	sebenernyo	bener	bener	Benar	Ya	Ada
551	sebij	biji	biji	Benar	Ya	Ada
552	seburuk	buruk	buruk	Benar	Ya	Ada
553	secaro	secaro	secaro	Benar	Ya	Ada
554	sedengke	sedeng	sedeng	Benar	Ya	Ada



# Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
555	sedi	sedi	sedi	Benar	Ya	Ada
556	sedikitlah	sedikit	sedikit	Benar	Ya	Ada
557	<b>sedila</b>	<b>la</b>	<b>sedi</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
558	segalo	galo	galo	Benar	Ya	Ada
559	sejak	sejak	sejak	Benar	Ya	Ada
560	sejalan	jalan	jalan	Benar	Ya	Ada
561	sejek	sejek	sejek	Benar	Ya	Ada
562	sekali	sekali	sekali	Benar	Ya	Ada
563	sekarang	karang	karang	Benar	Ya	Ada
564	sekolah	sekolah	sekolah	Benar	Ya	Ada
565	selamat	lamat	lamat	Benar	Ya	Ada
566	selamo	selamo	selamo	Benar	Ya	Ada
567	semu	semu	semu	Benar	Ya	Ada
568	<b>semuanyo</b>	<b>bua</b>	<b>semua</b>	<b>Benar</b>	<b>Tidak</b>	<b>Ada</b>
569	senang	senang	senang	Benar	Ya	Ada
570	sepanjang	panjang	panjang	Benar	Ya	Ada
571	sesampi	sampi	sampi	Benar	Ya	Ada
572	sesuda	suda	suda	Benar	Ya	Ada
573	setuju	setuju	setuju	Benar	Ya	Ada
574	sewaktu	waktu	waktu	Benar	Ya	Ada
575	siapo	siapo	siapo	Benar	Ya	Ada
576	sikok	sikok	sikok	Benar	Ya	Ada
577	soal	soal	soal	Benar	Ya	Ada
578	sopan	sopan	sopan	Benar	Ya	Ada
579	sopir	sopir	sopir	Benar	Ya	Ada
580	suaro	suaro	suaro	Benar	Ya	Ada
581	subur	subur	subur	Benar	Ya	Ada
582	suda	suda	suda	Benar	Ya	Ada
583	sudah	sudah	sudah	Benar	Ya	Ada
584	sudahlah	sudah	sudah	Benar	Ya	Ada
585	sudo	sudo	sudo	Benar	Ya	Ada
586	sukonyo	suko	suko	Benar	Ya	Ada
587	sungai	sungai	sungai	Benar	Ya	Ada
588	sungi	sungi	sungi	Benar	Ya	Ada
589	supayo	supayo	supayo	Benar	Ya	Ada
590	tadi	tadi	tadi	Benar	Ya	Ada
591	tahan	tahan	tahan	Benar	Ya	Ada
592	tamba	tamba	tamba	Benar	Ya	Ada
593	tanah	tanah	tanah	Benar	Ya	Ada
594	tanahnya	tanah	tanah	Benar	Ya	Ada
595	taneman	tanem	tanem	Benar	Ya	Ada
596	tangkupla	tangkup	tangkup	Benar	Ya	Ada
597	tanyoke	tanyo	tanyo	Benar	Ya	Ada

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
598	tapi	tapi	tapi	Benar	Ya	Ada
599	tau	tau	tau	Benar	Ya	Ada
600	tawaan	tawa	tawa	Benar	Ya	Ada
601	tebela	bela	bela	Benar	Ya	Ada
602	tedampar	dampar	dampar	Benar	Ya	Ada
603	tedamparnya	dampar	dampar	Benar	Ya	Ada
604	tedenger	denger	denger	Benar	Ya	Ada
605	teguling	guling	guling	Benar	Ya	Ada
606	tekejut	kejut	kejut	Benar	Ya	Ada
607	tempat	tempat	tempat	Benar	Ya	Ada
608	tenagonyo	nago	tenago	Benar	Tidak	Ada
609	tenga	tenga	tenga	Benar	Ya	Ada
610	terang	terang	terang	Benar	Ya	Ada
611	teraso	raso	raso	Benar	Ya	Ada
612	terimo	terimo	terimo	Benar	Ya	Ada
613	terus	terus	terus	Benar	Ya	Ada
614	tetahan	tahan	tahan	Benar	Ya	Ada
615	tiang	tiang	tiang	Benar	Ya	Ada
616	tiap	tiap	tiap	Benar	Ya	Ada
617	tidak	idak	idak	Benar	Ya	Ada
618	tigo	tigo	tigo	Benar	Ya	Ada
619	tinggal	tinggal	tinggal	Benar	Ya	Ada
620	tinggalan	tinggal	tinggal	Benar	Ya	Ada
621	tinggalke	tinggal	tinggal	Benar	Ya	Ada
622	tinggalla	tinggal	tinggal	Benar	Ya	Ada
623	tolong	tolong	tolong	Benar	Ya	Ada
624	tolongla	tolong	tolong	Benar	Ya	Ada
625	tolonglah	tolong	tolong	Benar	Ya	Ada
626	Tuhan	Tuhan	Tuhan	Benar	Ya	Ada
627	tulung	tulung	tulung	Benar	Ya	Ada
628	tumbu	tumbu	tumbu	Benar	Ya	Ada
629	tunggu	tunggu	tunggu	Benar	Ya	Ada
630	tuo	tuo	tuo	Benar	Ya	Ada
631	turun	turun	turun	Benar	Ya	Ada
632	ujan	ujan	ujan	Benar	Ya	Ada
633	ujinyo	uji	uji	Benar	Ya	Ada
634	untalke	untal	untal	Benar	Ya	Ada
635	untuk	untuk	untuk	Benar	Ya	Ada
636	untung	untung	untung	Benar	Ya	Ada
637	urusan	urus	urus	Benar	Ya	Ada
638	usul	usul	usul	Benar	Ya	Ada
639	usulke	usul	usul	Benar	Ya	Ada
640	wakil	wakil	wakil	Benar	Ya	Ada





### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

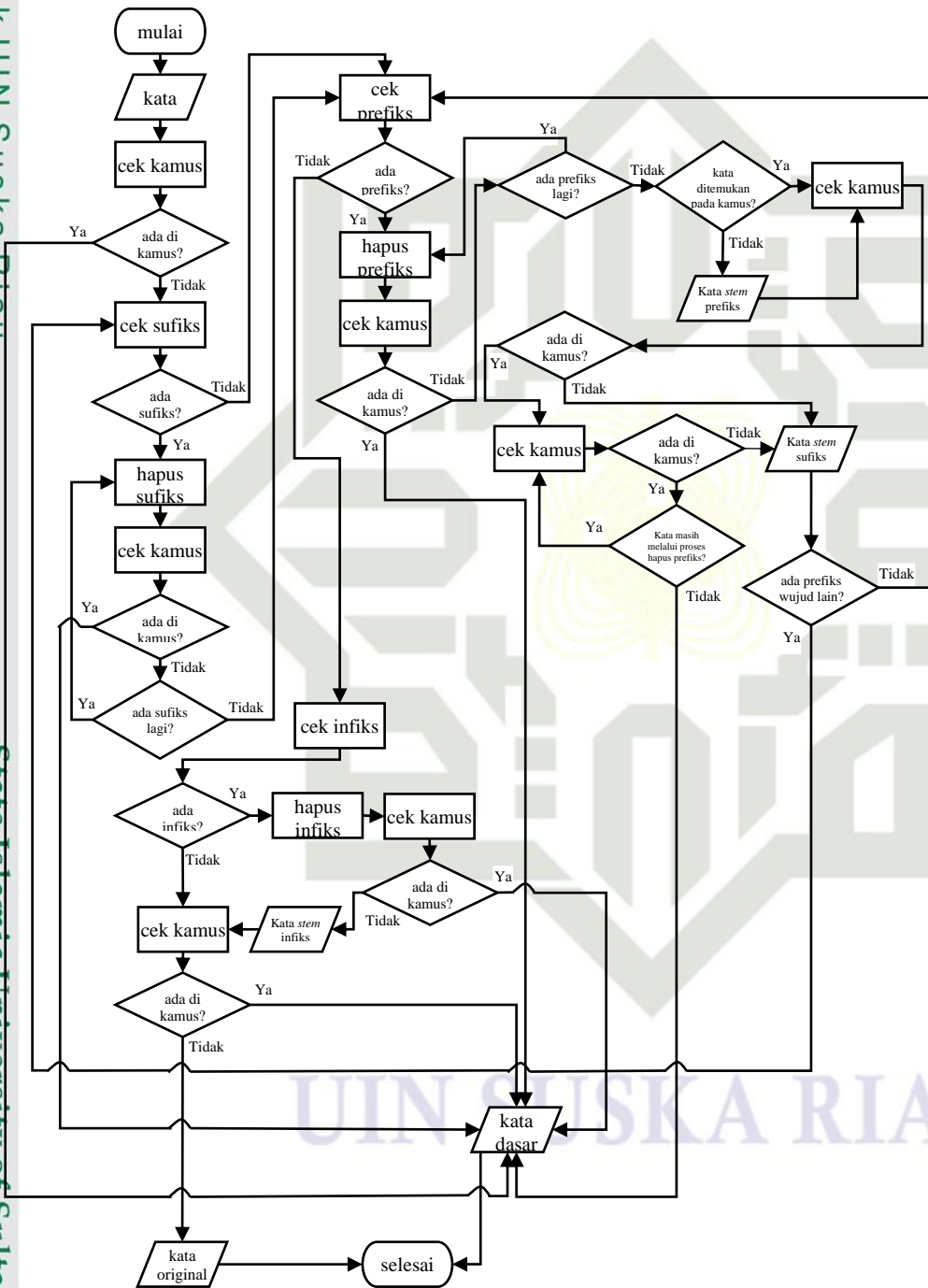
No	Kata Uji	Stemming Algoritma	Stemming Validator	Hasil Sistem	Sesuai Harapan	Kamus
641	waktu	waktu	waktu	Benar	Ya	Ada
642	warga	warga	warga	Benar	Ya	Ada
643	wong	wong	wong	Benar	Ya	Ada
644	yang	yang	yang	Benar	Ya	Ada
645	zaman	zaman	zaman	Benar	Ya	Ada



UIN SUSKA RIAU

## LAMPIRAN B

### FLOWCHART ALGORITMA STEMMING BASO PELEMBANG (KOMBINASI 2)



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

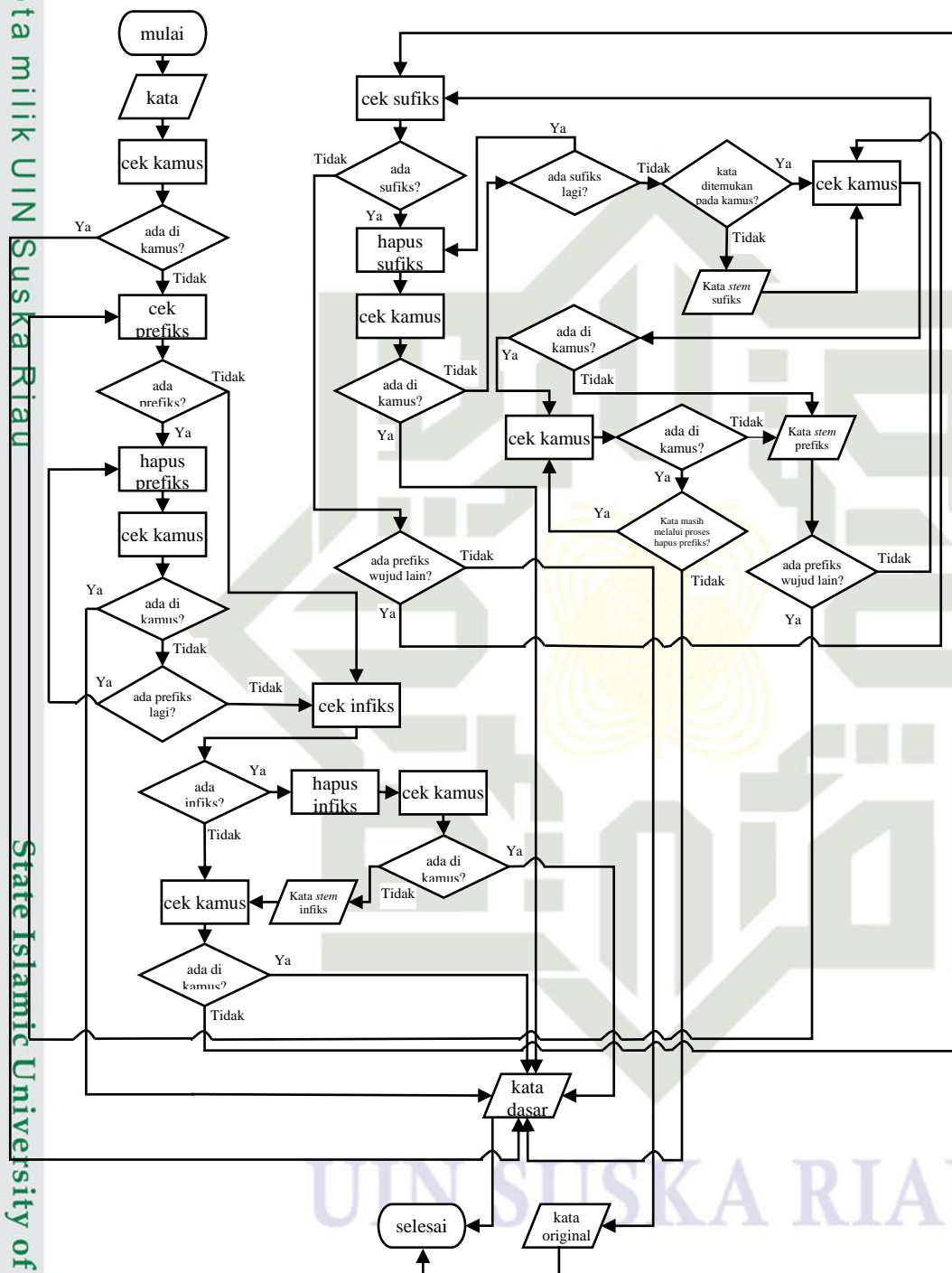
# Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## FLOWCHART ALGORITMA STEMMING BASO PELEMBANG (KOMBINASI 3)





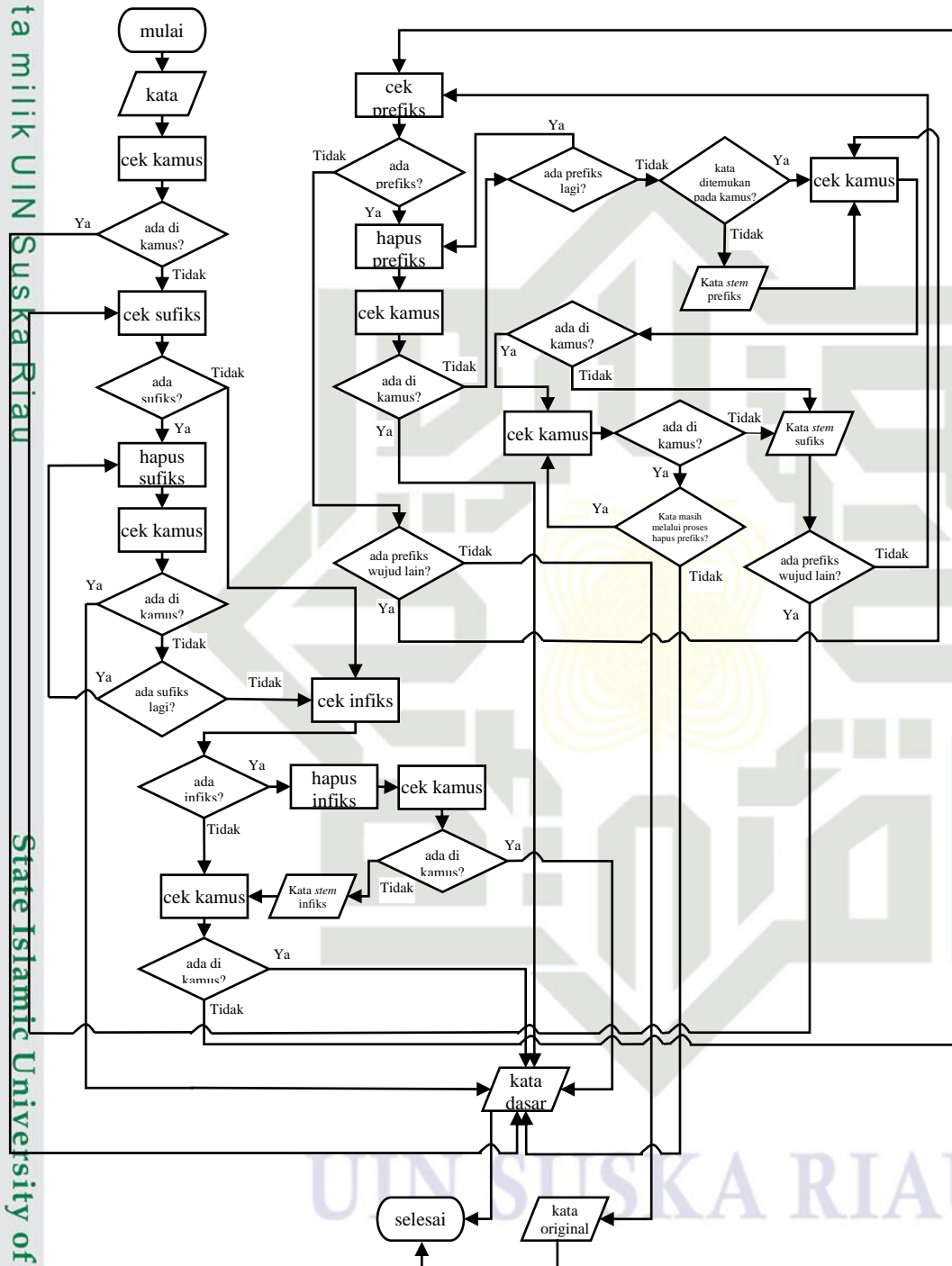
### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

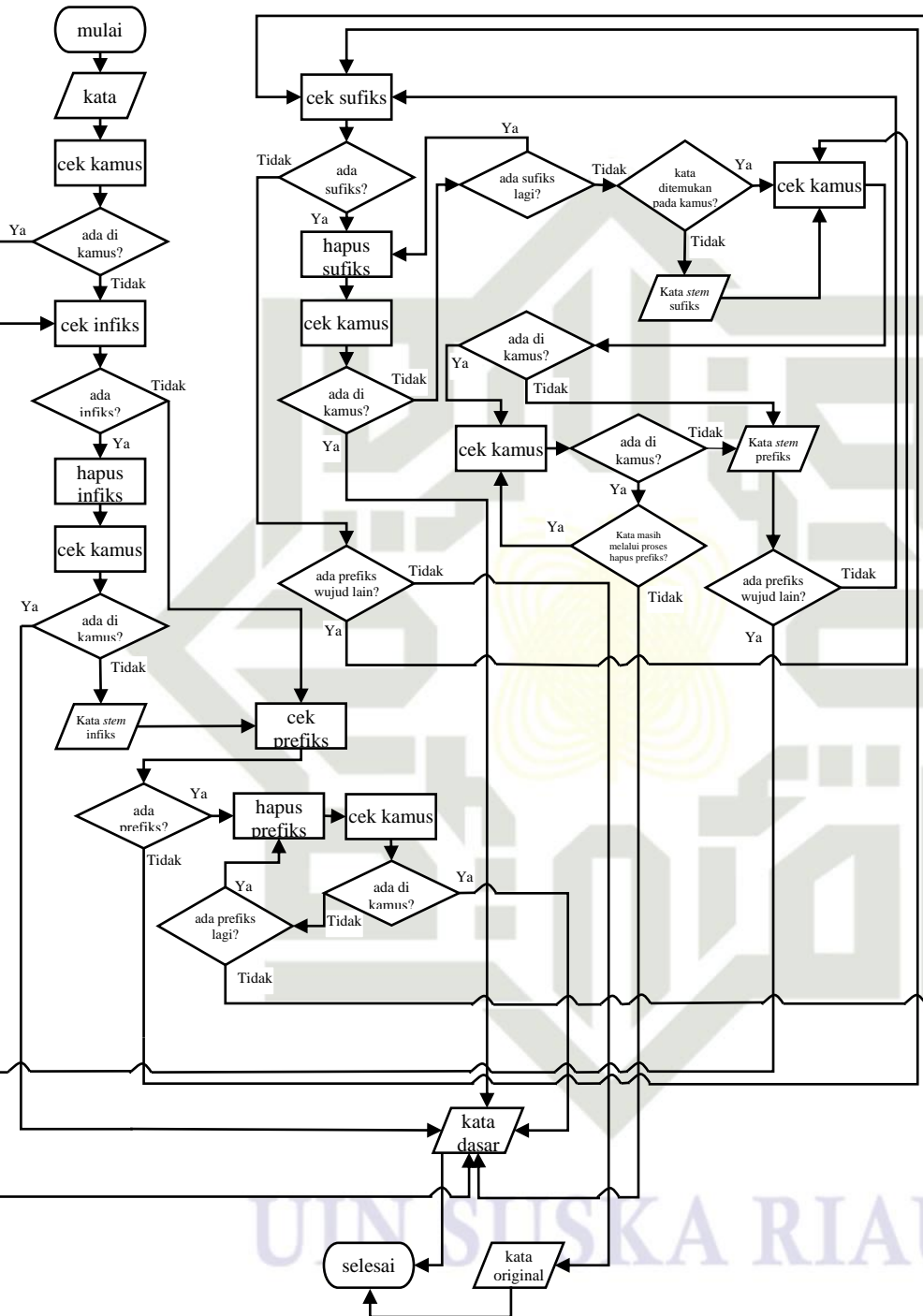
- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

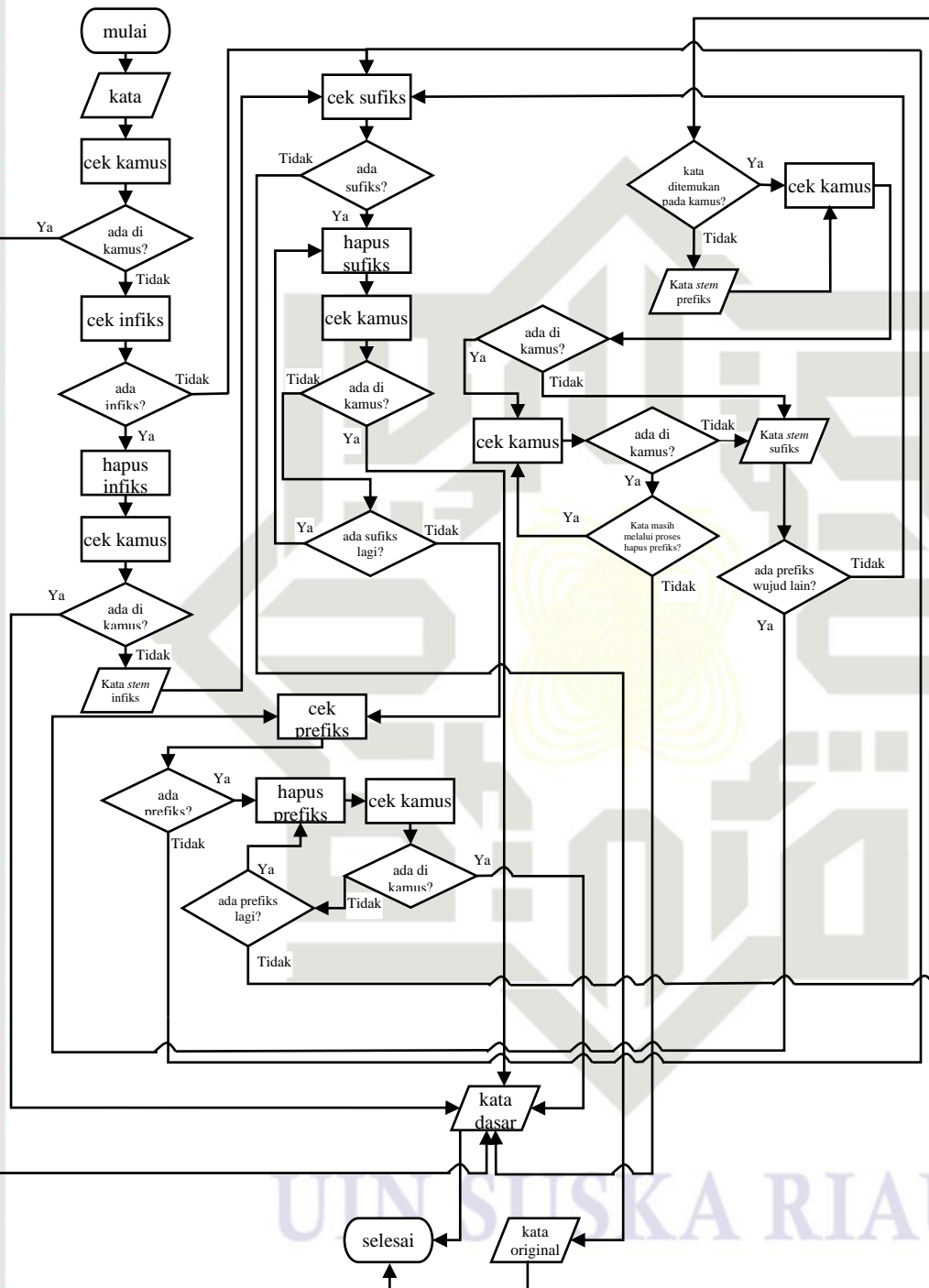
## FLOWCHART ALGORITMA STEMMING BASO PELEMBANG (KOMBINASI 4)



**FLOWCHART ALGORITMA STEMMING BASO PELEMBANG  
(KOMBINASI 5)**



## FLOWCHART ALGORITMA STEMMING BASO PELEMBANG (KOMBINASI 6)



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**LAMPIRAN C****LEMBAR VALIDASI DATA STEMMING****LEMBAR VALIDASI DATA STEMMING**

Yang bertanda tangan dibawah ini:

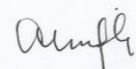
Nama : R. M. ALI HANAFIAH

Usia : 60 Tahun

Pekerjaan : Budayawan Kota Palembang

Menyatakan bahwa hasil data *stemming* manual pada penelitian yang berjudul "ALGORITMA STEMMING UNTUK BASO PELEMBANG BERBASIS ATURAN TATA BAHASA" telah sesuai dengan kata dasar yang sebenarnya dari *Baso Palembang*. Demikian lembar pengesahan ini dibuat untuk digunakan dalam melengkapi laporan tugas akhir.

Palembang, ..... September 2019

  
R.M. Ali Hanafiah  
(Validator)

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN D

### BIODATA VALIDATOR



Nama Lengkap	: RM Ali Hanafiah
Nama Panggilan	: Mang Amin
Tempat, Tanggal Lahir	: Palembang, 23-09-1959
Jenis Kelamin	: Laki-laki
Alamat	: Jl. Ki Gede Ing Suro No.16 RT.012 / RW.003, Kel/Desa. 28 Ilir, Kec. Ilir Barat II, Kab/Kota. Palembang, Sumatera Selatan
Kontak/WhatsApp	: +62852-7345-7889
Agama	: Islam
Status Perkawinan	: Cerai Hidup
Pekerjaan	: Pegawai Negeri Sipil dan Budayawan
Kewarganegaraan	: WNI
Golongan Darah	: O



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN E

### DAFTAR RIWAYAT HIDUP

#### FINA DEVIANA

finadeviana@students.uin-suska.ac.id | +6281266311157

Tempat Tanggal Lahir	Kuningan, 06 Mei 1997
Alamat	Jl. Harapan Raya No.50
Genus Kelamin	Perempuan
Agama	Islam
Tinggi Badan	172 cm
Kewarganegaraan	WNI
Motto	Sabar, Ikhlas, dan Tawakkal



#### PENDIDIKAN

Tahun 2002-2003	<b>TK Ade Irma Suryani Nasution - Kuningan</b>
Tahun 2003-2004	<b>SD Negeri Bojong 1 (Kelas 1) - Kuningan</b>
Tahun 2004-2009	<b>SD Negeri Mekarsari 09 (Kelas 2 hingga Tamat) - Bekasi</b>
Tahun 2009-2012	<b>SMP Negeri 1 Tambun Selatan - Bekasi</b>
Tahun 2012-2015	<b>SMA Swasta PGRI Pekanbaru - Pekanbaru</b>
Tahun 2015-2019	<b>S1 Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau</b>

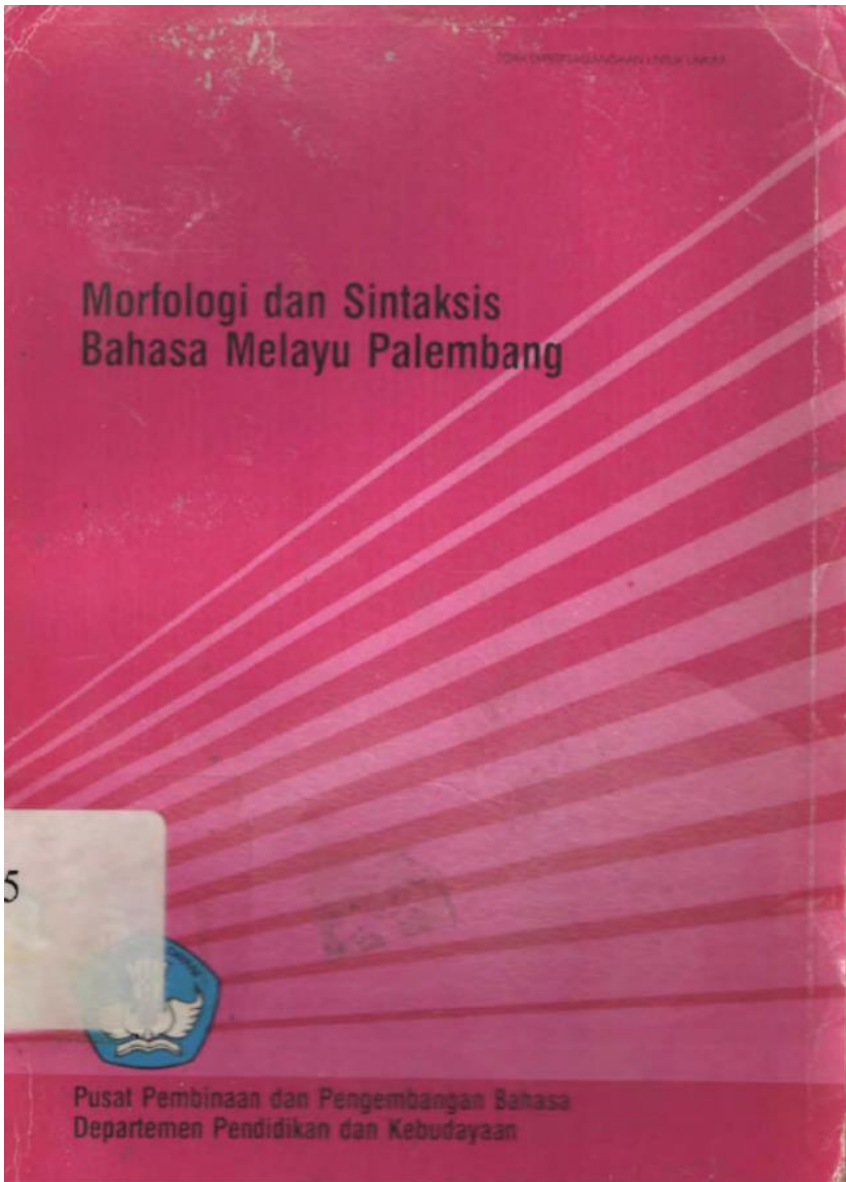


**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN F

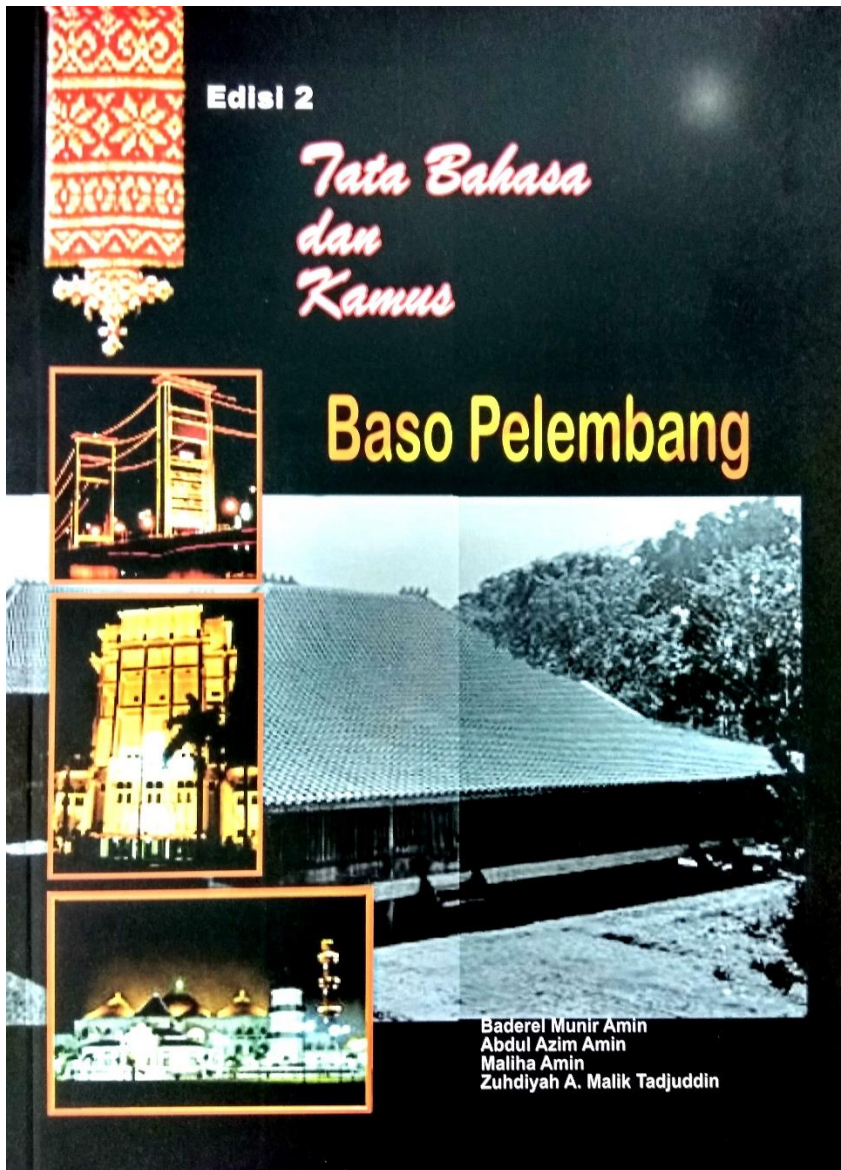
### BUKU MORFOLOGI DAN SINTAKSIS BAHASA MELAYU PALEMBANG (1987)



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**BUKU TATA BAHASA DAN KAMUS *BASO PELEMBANG* (2009)**







## NOVEL *JUARO* (2005)



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.